



Transcriptional Regulation in Eukaryotes
Concepts, Strategies, and Techniques
Second Edition

真核生物转录调控
概念、策略与技术
（第二版）

〔美〕M. F. 凯里　C. L. 彼得森　S. T. 斯梅尔　著
王进科　主译

科学出版社

生命科学实验指南系列·典藏版

真核生物转录调控

概念、策略与技术

（第二版）

Transcriptional Regulation in Eukaryotes

Concepts, Strategies, and Techniques

Second Edition

〔美〕M. F.、凯里　C. L. 彼得森　S. T. 斯梅尔　著

王进科　主译

科学出版社

北京

图字：01-2011-3184号

内 容 简 介

“生命科学实验指南系列”图书均出自名家，包括众多从 Cold Spring Harbor Laboratory Press 和 John Wiley & Sons 等国际知名出版社引进的实验室必备工具书，是生命科学领域最先进、实用、权威的实验手册类优秀图书。该系列图书简单明了，囊括了全世界最著名的生物类实验室操作方法，无论是初学者还是需要深入研究的科研工作者都能从中获益。该系列图书在读者群中有较高的知名度和美誉度，特别是以《分子克隆实验指南》和《精编分子生物学实验指南》为代表，堪称经典，分别被喻为生命科学领域的“蓝宝书”和“红宝书”。现挑选其中的精品集结成典藏版。

图书在版编目（CIP）数据

生命科学实验指南系列：典藏版/雷东锋等编著. —北京：科学出版社，2016
ISBN 978-7-03-047486-5

Ⅰ.①生… Ⅱ.①雷… Ⅲ.①生命科学–实验–指南 Ⅳ.①Q1-0

中国版本图书馆 CIP 数据核字(2016)第 043878 号

责任编辑：王 静 李 悦
责任印制：张 伟 / 封面设计：刘新新

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

北京厚诚则铭印刷科技有限公司 印刷
科学出版社发行 各地新华书店经销
*
2016年7月第 一 版 开本：787×1092 1/16
2016年7月第一次印刷 印张：1310 1/2
字数：31 074 000

定价：4500.00 元

(如有印装质量问题，我社负责调换)

译 者 序

基因组学的研究表明，即使像人类这样高等而复杂的生物，其基因组所蕴含的基因只不过3万多个，而与人类看起来截然不同的物种大猩猩，竟然基因组成基本和人类相同，因此，人类与大猩猩的差异并不取决于基因的不同，而在于一套相同基因的不同调控，人与人之间的差异更是如此。可见，基因调控很重要，调控不同，产生的结果可能相去万里。

生物在演化的过程中，逐渐形成了复杂的调控体系，这种调控体系极其严密，从其调控层次的多样性就可见一斑。这些层次包括转录调控、RNA加工调控、翻译调控等。除了这些经典的调控途径外，现在已经知道细胞中的染色质结构及状态、无处不在的大量非编码RNA、表观遗传，甚至以前视为垃圾的大量非编码DNA，都已经出现在基因调控的武器库中，仿佛细胞核中的一切都是为了基因的表达调控而存在。

当然，在这些层次中，转录调控居于非常枢纽的位置。因此，关于转录调控，特别是真核生物的转录调控，有海量的研究论文和不少的专著。在这些璀璨的文献中，就有美国加利福尼亚大学Michael F. Carey等编著的《真核生物转录调控——概念、策略与技术》一书。正如其名，该书第一版于2000年出版时，就系统地阐述了真核生物转录调控的基本概念、研究策略和实验技术。

而自那以后的10年，恰好是生命科学研究技术发生巨大变化的10年，在此期间涌现了对研究基因表达调控非常重要的许多生物的完整基因组序列，诞生了重量级的技术武器，如染色质免疫沉淀、RNA干扰、DNA微阵列芯片、先进的DNA测序技术、蛋白质组学和生物信息学等技术。这些技术的应用产生了前所未有的基因调控信息。为了充分反映这些成果，Michael F. Carey，Craig L. Peterson和Stephen T. Smale对第一版进行了彻底的重写，并增加了重要的新内容（如第13章），于2009年出版了该书的第二版，即由我们此次翻译的版本。

我们翻译该书出于三点原因，一是该书是由大名鼎鼎的美国冷泉港实验室出版社出版的关于真核基因转录调控的经典专著，该书非常有助于相关人员掌握基础理论知识和实验操作实践，我们是该书第一版的忠实读者及其策略和技术的使用者；二是我们的实验室在近10年来执著地追求和热爱真核基因转录调控领域的研究；三是我们希望为国内同行做一点贡献，为他们的学习、研究提供一点帮助和便利，以促进国内在该领域的人才培养、教学和科研。

翻译该书的人员基本上都是东南大学生物电子学国家重点实验室王进科课题组的研究生及教学科研老师。在此，对参与本书翻译的研究生、博士后致以诚挚的感谢，还要特别感谢白云飞、李同祥两位老师，他们应邀翻译了部分内容。

本书的翻译分别由以下人员完成：邢宇俊（第1章），郭伟（第2章），李同祥（徐州工学院）（第3章），杨洋、华雯妤（第4章），谷光明（第5章），潘艳（第6章），

陈鸿飞（第7章），高婧（第8章），陈莹、凌小倩（第9章），刘颖勋（第10章），白云飞、韩芹（第11章），王进科（前言、概述、缩略词、第12章、第13章、附录、索引）。为了保持全书译文的准确性、前后内容的连贯性，以及翻译风格的一致性，王进科对全书各章节的翻译进行了逐字逐句的修改和润色，完成了终稿。

科学出版社的夏梁编辑为本书翻译的立项、版权的获取及出版给予了大力支持，在此致以由衷的感谢。我们还要特别感谢国家重大科技专项（转基因：2009ZX08008－007）和国家自然科学基金项目（60871014）对该书翻译和出版的资助。

最后，由于译者能力所限，翻译中若有不妥之处，敬请读者见谅，同时欢迎读者们批评指正。

王进科

2011年10月

前　言

当这本书的第一版于2000年出版时，真核细胞转录调控的分子和生化方面已经进行了20多年相当深入的研究。那段时间已经出现的实验策略，不断成功地取得了关于个别基因调控机制或个别转录因子功能调控机制的宝贵见解。但同时，每种策略常见的缺陷和局限性也显露无遗。虽然许多技术的实验方案触手可及，但逐步完成一项分析所必需的策略决策却始终没有被归纳总结。这方面的不足对我们两人（M. F. C. 和 S. T. S.）尤为明显，因为当时我们作为讲师，指导冷泉港实验室每年夏天开设的真核基因表达课程。这个实验室课程是专为下列人员量身定做的，包括对理解一个特定疾病相关基因调控感兴趣的医生、在其他领域训练有素的对一个特定生物过程相关基因调控机制感兴趣的科学家，以及启动转录调控项目的研究生或博士后研究人员。这本书的第一版针对这些不同群体的科学家。我们聚焦于普通研究人员在进行转录调控分析时所面对的问题，并概述了完成一项分析所推荐的策略。

自第一版出版以来的9年中，转录调控领域已经见证了巨大的变化，宛如20世纪70年代分子克隆技术诞生一样璀璨。在此期间取得了许多进展，这些进展极大地促进了对个别基因调控机制和个别转录因子功能的分析。其中，最引人注目的是发展了用于监测内源性座位上蛋白质-DNA相互作用的强大的染色质免疫沉淀（ChIP）分析，用于进行转录调控因子功能缺失研究的RNA干扰（RNAi）策略，以及用于检测多蛋白复合物和翻译后修饰的先进的质谱分析方法。

意义更大的是，目前已经获得了越来越多的物种的完整基因组序列，使研究人员以最小的努力分离基因及其调控区域，进行复杂的系统发生分析以鉴定进化上保守的区域。也出现了利用基因组序列信息巨大财富的强大的实验工具，包括监测基因表达的微阵列，以及检测全基因组蛋白质-DNA相互作用的ChIP-chip和ChIP-Seq方法。

在概念层次上，染色质结构和组蛋白修饰在基因调控中的重要的和动态的作用，在20世纪90年代后5年中才刚刚开始出现，但现在却已广受重视。对细胞核组织和染色体动态学研究的重要性也日益受到重视，因为它们和谐地安排着特定组合基因中协调的差异的调控，而这些基因组合决定了生物的发育或刺激应答。

总体来说，这些概念和技术方面的进步对转录调控领域极为有利，并大大地增加了该领域的复杂性和作出重要进展的速度。然而，在全基因组方法中取得的海量信息，以及参与基因调控的大量的染色质重塑机器和修饰酶的详尽知识，都给试图破译机制的研究人员带来沉重的负担。第二版在如何研究个别基因的调控及其转录因子的架构中，考虑到了这些新的进展。对此版本的一个主要改变是第三个作者（C. L. P.）的参与，他带来了在染色质前沿领域多年研究的经验。本书已经对前面的章节进行了重新写作，其中，对染色质给予特别的重视。此外，增加了一个致力于染色质生物化学的章节，以及一个以核小体环境中的转录调控为重点的全新的和更新的前言章节（第1章）。由于不

断增加的复杂性和该领域的多样性，不可能全面地描述可能对转录调控研究者有价值的每个实验策略和技术。但是，我们试图总结对初进转录领域及已身处该领域而要解决调控机制的一个新问题的研究人员最感兴趣和最相关的概念、方法和技巧。

我们再次感谢对本书第一版做出重要贡献的许多研究者，包括 Doug Black、Ken Burtis、Grace Gill、David Gilmour、Jim Goodrich、Steve Hahn、Mike Haykinson、Leila Hebshi、Reid Johnson、Marc Learned、Ranjan Sen、Bill Tansey 和 Amy Weinmann。我们也要深深感谢很多为第二版作出贡献的同事，包括 Dev Bhatt、Josh Black、Craig Crockett、Justin Lin、Scott Pope、Michael Zhang，特别是 John Heiss，他的结构渲染技能深切地反映在封面插图上。我们对 Judy Cuddihy 的援助表示感谢，她是我们在冷泉港实验室出版社的编辑。此外，我们也感谢冷泉港实验室出版社的许多人的重要贡献，包括 Kathleen Bubbeo、Mary Cozza、Jan Argentine、Denise Weiss、Martin Winer、Lauren Heller 和 Alex Gann。最后，我们感谢冷泉港实验室出版社的发行人 John Inglis，他的支持对这个新项目的完成至关重要。

M. F. 凯里
C. L. 彼得森
S. T. 斯梅尔

概　述

本书的目的是，当为新分离的基因进行转录调控机制的分子分析，或为新的转录因子或染色质复合物进行生化分析时，提供所用方法及应考虑问题的详尽描述。我们的重点是哺乳动物转录，因为调控的组合特性和简便遗传学的缺乏使其复杂化。我们不时地提及酵母、果蝇和其他生物中的研究，因为对这些生物的研究具有更容易操控的遗传学方法，使得对特定机制问题有了详尽的认识。书中涵盖的论题，从确定一个基因是否事实上在转录起始水平上被调控，一直延伸到用于表征染色质环境中转录调控的生化机制的先进策略。尽管包含了很多专门的和详尽的技术，但本书的独特特征在于它在策略和概念上强调单个基因和转录因子的分析，以及调控它们的染色质事件。

第 1 章综述了 RNA 聚合酶 II 转录领域的现状。本章在第二版中已经整体改写，其中高度重视从传统的生化和分子方法，以及当代基因组学方法获得的有关机制的知识。本章为全书引言，目的是为进入本领域的研究者提供一个全面的综览，使他们对本领域活跃研究的范畴及将要面对的调控策略的种类有所了解。本章中我们界定了已知调控区域、转录机器部件、激活因子及抑制因子的一般属性。我们还广泛地讨论了染色质的结构、修饰和重塑，强调了基因激活和抑制。本章还为对各种论题的其他细节感兴趣的新手收录了精选的综述文章。

第 2 章介绍了在开始分析一个目的基因的调控机制或一个新转录因子的功能和靶标之前，需要考虑的一般策略问题。我们首先讨论了分析的目标，本章纳入这个论题的原因是许多研究者带着不切实际的期望进入转录领域。据推测，之所以出现这些期望，是因为使用基本报告分析、电泳迁移率变动分析（EMSA），或全基因组染色质免疫沉淀（ChIP）实验，进行一个控制区或转录因子的初步分析是相对简单的，而要在了解基因调控机制或转录因子功能方面取得有意义的进展，通常需要做出大量的努力。第 2 章还讨论了实现这些目标的可行性，这种可行性在很大程度上取决于是否能够获得特定的工具，包括用于功能和生化研究的适当的细胞系，以及恰当的功能分析方法。

一个在第 2 章和所有随后的章节中将变得显而易见的问题是，本书并没有给大部分所描述的方法提供详尽的方案，相反，对标准方法、手册和其他发表的方案仅给出了参考文献，目的是避免重复现有文献中提供的有价值的信息，从而将重点放在其他地方找不到的策略咨询上。虽然我们可以将本书写得没有任何方案（因为大部分方案可在文献中找到），但我们还是包括了一些精选的方案，原因有三点。首先，有些方案被选入，是因为我们感到读者会从具体步骤的详细解释和方法学的历史中受益。其次，在某些情况下，当读者阅读本书时，我们感到有必要给读者提供一种技术的过程感。最后，有几个方案被收录是因为它们的特殊性质和这些方法事实上没有常规的出处。

第 3 章～第 8 章设想为那些愿意为一个已鉴定的新基因探索其调控机制的研究者提供逐步的指导。

如果研究者选择通过研究基因的启动子开始分析，在第 3 章和第 4 章中首先讨论了重要的考虑。第 3 章介绍了确定转录起始位点所在位置的方法，这是每个启动子分析中重要的第一步。自本书第一版出版以来，在确定转录起始位点方面，取得了方法学上的重大改进，包括快速扩增 cDNA 5′ 端（5′RACE）方法的建立，该方法强烈倾向于全长加帽转录物，并且最大限度地减少了多年来困扰起始位点作图的背景问题。通过与鉴定转录起始位点的常规方法的比较，本章讨论了这些方法的优点和缺陷。此外，我们描述了使用这些新方法来鉴定基因组内所有基因转录起始位点的研究过程中所作的努力。

第 4 章描述了启动子功能分析方法的发展，也就是说，发展可用于监测启动子活性的分析方法，以便随后的启动子突变体和候选转录调控因子的研究。本章详细讨论了瞬时转染和稳定转染分析方法，包括转染程序、报告基因、载体和分析的概述，以及实验的最初设计和解释。本章还简要介绍了其他功能分析方法，包括体外转录和转基因分析，以及它们的优缺点。

第 5 章描写了鉴定一个目的基因远程控制区，如增强子、沉默子、基因座控制区，以及边界元件的方法。在本书第一版中，远程控制区的鉴定被描述为一个艰难的挑战，以 DNase I 高敏感性和功能性方法为实现这一目标的首选方法。然而，大量物种的完整基因组序列的获得，以及利用这些序列信息财富的全基因组生物信息学技术和实验技术的出现，已对鉴定重要控制区产生重大影响。虽然传统的 DNase I 高敏感性和功能性方法仍然具有相当大的价值，但来自大量相关物种的基因座的系统发生比较有不可估量的价值。用于鉴定 DNase I 高敏感性位点或与增强子、沉默子及边界元件已知相关的组蛋白修饰富集区的全基因组方法，为候选调控区的鉴定提供了其他选项。从功能的角度来看，细菌人工染色体转基因已经成为鉴定对目的基因恰当调控重要的 DNA 区域的常见而有价值的工具。

分离出启动子或远程控制区后，下一步就是要鉴定负责控制区功能的个别 DNA 模体，这一步将在第 6 章中进行描述。此外，已经证明许多物种基因组序列的获得价值巨大，这使得鉴定进化上保守的 DNA 模体成为可能。自本书第一版出版后，另一主要进展是发展了强大的 ChIP 分析，用于监测候选转录调控因子对内源性控制区的结合。用以揭示已知转录因子在 DNA 片段及整个基因组中的潜在识别序列的生物信息学程序也得到了改善。在常规诱变策略的背景中，本章讨论了这些进展的影响；对于评价一个目的 DNA 模体的功能相关性，以及鉴定被生物信息学和 ChIP 方法遗失的重要模体而言，常规诱变策略依然重要。

当重要的 DNA 元件被鉴定，但其相关结合蛋白不太明确时，就必须鉴定候选结合蛋白。第 7 章描述了实现这一目标的策略，并强调了日益先进的质谱方法在鉴定可与 DNA 模体结合的蛋白质及蛋白质复合物中的价值。本章还介绍了检测核提取物中蛋白质-DNA 相互作用的常规方法，尤其是 EMSA 和 DNase I 足迹，同时介绍了用于鉴定可与模体结合的转录因子的其他策略，包括酵母单杂交筛选、体外表达文库筛选，以及哺乳动物表达克隆。

通过专注于一个重要问题，即在鉴定了与一个重要 DNA 元件体外结合的因子后，如何才能确定该因子是否的确负责控制元件在体内的功能？第 8 章逐步概述了一个新基

因的表征。虽然没有实验可以提供确凿的证据，但 ChIP 分析和 RNA 干扰（RNAi）已经成为验证候选因子是基因重要而直接的调控因子的假设的强大工具。本章对这两种方法及其他一些实验策略进行了描述，同时对每种策略提供的信息及其局限性进行了讨论。

第 9 章描述了用于染色质体内表征的实验策略。作为监测特异性转录因子与基因组位点结合，以及修饰组蛋白在基因座或基因组内分布的方法，首先讨论了 ChIP、Dam-ID，以及使用 DNase I、硫酸二甲酯和高锰酸盐的基因组足迹法。其次，我们描述了几个基于微球菌核酸酶的监测核小体在基因座或基因组内存在及定位的策略。本章涵盖了用于评估基因激活和沉默期间染色质结构动态变化的核酸酶可及性分析，还讨论了用于检验控制区之间染色质内及染色质间相互作用的染色质构象捕获分析。最后，本章简要讨论了监测个别控制区及全基因组水平 DNA 甲基化的传统的和最近的策略。

第 10 章～第 13 章从生化角度强调了基因调控的机制。如何划定序列特异性转录因子或共激活因子/抑制因子的结构域？激活因子和抑制因子如何将共激活因子和共抑制因子募集到启动子上，以及它们在重组转录系统中的功能效应是什么？如何研究染色质修饰和重塑复合物的结合和活性？

一般情况下，如果研究者要通过一个分析的不同阶段，划定参与与其他调控蛋白和转录机器相互作用的蛋白质结构域就成了当务之急。该信息对于完成机制的生化分析至关重要。用来获取这种深刻见解的方法被称为“结构-功能”分析。这不是一个简单的任务，并且方法和决策往往基于所研究的调控蛋白的特定类型。第 10 章从几个方面讨论了结构-功能分析。本章简述了用于研究蛋白质相互作用的方法，以便让研究者设计特定的分析方法分析有关结构域。本章讨论了一种简单的缺失分析，作为划定调控蛋白的不同区域怎样对 DNA 结合和转录调控的不同方面发挥作用的手段。将这一讨论作为跳板可跨向更先进的方法，包括结构域交换——一个把精确的功能归于蛋白质组成部件的简单手段。但是，最重要的是，转录的分子认识往往来源于对介导相关接触的特定氨基酸残基的了解。特别的重点放在引导研究者通过不同概念的方法产生定点突变体、这些突变体如何建模，以及将突变发生与晶体结构的结果进行比较的案例研究上。不幸的是，这些类型的分析在该领域中尚未明确界定。为了提供某种水平的指导，本章使用了该领域的一些顶级实验室如何研究各种激活因子、抑制因子，以及共激活因子和共抑制因子的很多例子。

DNA 的蛋白质组合识别是转录应答的细胞和发育特异性的主要贡献者。无论是单独的还是与其他蛋白质协同，蛋白质用来结合启动子或增强子的机制，都是转录领域中研究的关键领域。随着新的转录因子被鉴定，基因组范围的研究开始进入一个更加机制性的阶段，更多的重点将放在了解 DNA-结合协同性和组合相互作用上，特别是共调控基因上的 DNA-结合协同性和组合相互作用。第 11 章描述了平衡结合的原理。本章介绍了 DNA 识别的概念，为新手描述了 DNA-蛋白质相互作用的化学方法，最后讨论了化学方法和核酸酶探针如何用于产生 DNA 结合的详细模型。此外，本章概述了将从化学探测中衍生出来的模型与 DNA-蛋白质共复合物的晶体结构结果进行比较的个案研究。最后但最重要的是，本章为被称为增强体的核蛋白复合物提供了概念和研究的基本

介绍，这是一个构成转录因子组合作用基础的研究领域。

最终，研究者可能想了解受激活因子影响的详细的生化步骤。这个目标包括两项任务：①发展体外再现调控现象的强大的体外转录系统；②以高度专业化的试剂，包括纯化的转录因子和染色质模板，设计原理性实验。第 12 章指导研究者通过必要的逻辑判断和试剂，设计恰当的报告模板，并发展活跃的转录系统。本章讨论了如何测定和优化体外转录反应，包括无 G 序列盒和引物延伸。本章还介绍了用粗制的或纯化的通用因子和聚合酶 II 产生重建系统的现有的方法。一旦体外转录系统准备就绪，研究者可能希望进一步使用纯化的转录试剂进行激活转录的生化机制研究。就新的概念和专门的试剂两个方面而言，这是一个迅速发展的领域。本章概述了很多策略，可以用来研究使用粗制或纯化试剂的激活转录中的特定步骤。其中包括用于分析转录复合物组装的方法，如固定模板方法、高锰酸盐探测及其他方法，重点是分析方法的发展和数据的解释。本章还试图提供一个专门试剂来源的最新表格，包括表达和纯化重组转录因子，以及诸如调解因子和 TFIID 的多组分复合物的系统。

要了解激活因子和抑制因子如何作用，就必须首先着眼于这些蛋白质如何影响染色质。本书第二版以全新的一章，即第 13 章结束，该章致力于描写对染色质修饰和重塑相关的蛋白质进行生化和机制分析时所用的最新方法。当试图在体外组装核小体以进行机制研究时，首先描述了重要的考虑，同时描述每个现有方法的优势和局限性。接下来，我们讨论了通常用来表征体外染色质组装质量和完整性的几种方法，其中强调每个方法所提供的信息。众所周知，组蛋白翻译后修饰对基因调控与染色质动力学是至关重要的，所以我们讨论了可用于在体外环境中修饰组蛋白的策略，然后强调了用于研究酶和多蛋白复合物的生化方法，这些酶和多蛋白复合物负责组蛋白的修饰、核小体重塑和 DNA 甲基化。转录领域的中心目标是使用核小体模板在无细胞分析中模仿转录的精确调控。最后，我们讨论了这方面的进展，以此结束本章和本书。

缩 略 词

除了公制测量（如毫升）和化学符号（如盐酸）的标准缩写，本手册中始终使用的缩略词如下。

3C，chromatin conformation capture，染色质构象捕获

ac，H-bond acceptor or acetyl，氢键受体或乙酰基
acetyl-CoA，acetyl-coenzyme A，乙酰辅酶 A
ACF，ATP-utilizing component of the Drosophila embryo extract nucleosome assembly system，果蝇胚胎提取物核小体组装系统的 ATP 利用组分
ACH，active chromatin hubs，活性染色质中心
AdMLP，adenovirus major late promoter，腺病毒主要晚期启动子
ADP，adenosine diphosphate，二磷酸腺苷
AR，androgen receptor，雄激素受体
ARID，AT-rich interaction domain，富含 AT 的互作结构域
ATF-2，activating transcription factor 2，激活转录因子 2
ATP，adenosine triphosphate，三磷酸腺苷

BAC，bacterial artificial chromosome，细菌人工染色体
BAF，brahma-related gene 1 associated factor，梵天相关基因 1 结合因子
Bbd，Barr body-deficient，巴氏体缺失
BDF1，bromodomain factor 1，溴域因子 1
β-gal，β-galactosidase，β-半乳糖苷酶
bHLH，basic helix-loop-helix，碱性螺旋-环-螺旋
BRD2，bromodomain protein 2，溴域因子 2
BRE，IIB recognition element，IIB 识别元件
Brgl，brahma-related gene 1，梵天相关基因 1
Brm 1，mammalian brahma 1，哺乳动物梵天 1
BSA，bovine serum albumin，牛血清白蛋白
bZIP，basic leucine zipper，碱性亮氨酸拉链

C/EBP，CCAAT enhancer-binding protein，CCAAT 增强子结合蛋白
CAGE，cap analysis gene expression，帽分析基因表达
cAMP，. cyclic AMP，环一磷酸腺苷

CAT，chloramphenicol acetyltransferase，氯霉素乙酰转移酶
CBP，CREB-binding protein，CREB 结合蛋白
ChIP，chromatin immunoprecipitation，染色质免疫沉淀
ChIP-chip，ChIP-on-chip，染色质免疫沉淀-芯片
ChIP-Seq，chromatin immunoprecipitation sequencing，染色质免疫沉淀-测序
CMV，cytomegalovirus，巨细胞病毒
CoA，coenzyme A，辅酶 A
CREB，cyclic AMP response element binding protein，cAMP 应答元件结合蛋白
Cryo-EM，cryo-electron microscopy，低温电子显微镜
CTCF，CCCTC-binding factor，CCCTC-结合因子
CTD，carboxy-terminal domain，羧基端结构域
CTs，chromosomal territories，染色体域

Dam，DNA adenine methyltransferase，DNA 腺嘌呤转甲基酶
DEAE，diethylaminoethanol，二乙氨基乙醇
DFO，deferoxamine，去铁敏
DHFR，dihydrofolate reductase，二氢还原酶
DMP，dimethyl pimelimidate，庚二亚氨酸二甲酯
DMR，differentially methylated region，差异甲基化区
DMS，dimethyl sulfate，硫酸二甲酯
DMSO，dim ethyl sulfoxide，二甲基亚砜
DNMTs，DNA methyltransferases，DNA 甲基转移酶
do，H-bond donor，氢键供体
DPE，downstream promoter element，下游启动子元件
DSIF，DRB-sensitivity inducing factor，DRB 敏感性诱导因子
DTT，dithiothreitol，二硫苏糖醇

EEO，electroendosmosis，电内渗
ELISA，enzyme-linked immunosorbent assay，酶联免疫吸附测定
EMSA，electrophoretic mobility-shift assay，电泳迁移率变动分析
ES cells，embryonic stem cell，胚胎干细胞
EST，expressed sequence tag，表达序列标签
ExoIII，exonuclease III，外切核酸酶 III

FACS，fluorescence-activated cell sorting，荧光激活细胞分选
FAD，flavin adenine dinucleotide，黄素腺嘌呤二核苷酸
FISH，fluorescent in situ hybridization，荧光原位杂交
FRET，fluorescence resonance energy transfer，荧光共振能量转移

GFP，green fluorescent protein，绿色荧光蛋白
GIS，genome identification signature，基因组鉴定签名
GPT，xanthine-guanine phospho ribosyltransferase，黄嘌呤-鸟嘌呤磷酸核糖基转移酶
GR，glucocorticoid receptor，糖皮质激素受体
GSC，genome signature cloning，基因组签名克隆
GST，glutathione-S transferase，谷胱甘肽-S转移酶
GTF，general transcription factor，通用转录因子

HA，influenza virus hemagglutinin，流感病毒血凝素
HAT，histone acetyltransferase，组蛋白乙酰转移酶
HBS，HEPES-buffered saline，HEPES-缓冲盐
HDAC，histone deacetylase，组蛋白去乙酰化酶
HEK，human embryonic kidney，人胚胎肾
HisD，histidinol dehydrogenase，组氨醇脱氢酶
HIV-1，human immuno deficiency virus type 1，1 型人类免疫缺陷病毒
HLH，helix-loop-helix，螺旋-环-螺旋
HMG，high-mobility group，高迁移率族
HMGA，high-mobility group A，高迁移率族 A
HMK，heart muscle kinase，心肌激酶
HMT，histone methyltransferase，组蛋白甲基转移酶
HPI，heterochromatin protein 1，异染色质蛋白 1
HPC，human polycomb protein，人多梳蛋白
HPLC，high-pressure liquid chromatography，高压液相色谱法
HS sites，hypersensitive sites，高敏感位点
HSF，heat shock factor，热激因子
HSV-1，herpes simplex virus type 1，1 型单纯疱疹病毒
HSV-TK，herpes simplex virus thymidine kinase，单纯疱疹病毒胸苷激酶
HTH，helix-turn-helix，螺旋-转角-螺旋
HTP，hydroxyapatite，羟基磷灰石

IFN-β，interferon-β，β-干扰素
Ig，immunoglobulin，免疫球蛋白
IL，interleukin，白细胞介素
Inr，initiator，起始子
IPTG，isopropyl-β-D-thiogalactopyranoside，异丙基-β-D-硫代半乳糖苷
IRF，interferon regulatory factor，干扰素调节因子

K_d，equilibrium dissociation constant，平衡解离常数
KID，kinase-inducible domain，激酶诱导域
KRAB，Kruppel-associated box，Kruppel 相关盒

LBD，ligand-binding domain，配体结合域
LCR，locus control region，基因座控制区
LEF-1，lymphoid enhancer binding factor 1，淋巴增强结合因子 1
LM-PCR，ligation-mediated polymerase chain reaction，连接介导的聚合酶链反应
lod，logarithm-of-the odds，几率对数
LPS，lipopolysaccharide，脂多糖
LSD1，lysine-specific demethylase，赖氨酸特异性脱甲基酶
LTR，long terminal repeat，长末端重复

MALDI-TOF/TOF，matrix-assisted laser desorption/ionization time of flight/time of flight，基质辅助激光解吸/电离飞行时间/飞行时间
MAR，matrix attachment region，基质附着区
MBD，methyl-binding domain，甲基结合域
MBT，malignant brain tumor，恶性脑肿瘤
MDC1，mediator of DNA damage checkpoint protein，DNA 损伤检查点蛋白介质 1
mDIP，methylated DNA immunoprecipitation，甲基化 DNA 免疫沉淀
Med，mediator，调解因子
MEFs，mouse embryonic fibroblasts，小鼠胚胎成纤维细胞
MLA，methyl lysine analog，甲基赖氨酸类似物
MMTV，mouse mammary tumor virus，小鼠乳腺肿瘤病毒
MNase，micrococcal nuclease，微球菌核酸酶
MPG，magnetic particle concentrator，磁性颗粒分离器
MPE，methidiumpropyl EDTA，溴乙非啶丙基 EDTA
MTE，motif ten element，模体十元件
MTX，methotrexate，氨甲蝶呤

N_3 RdUTP，5-［*N*-（*p*-azidobenzoyl）-3-aminoallyl］-dUTP，5-［*N*-（*p*-叠氮苯甲酰）-3-氨基烯丙基］-dUTP
NAD，nicotinamide adenine dinucleotide，烟酰胺腺嘌呤二核苷酸
NE，nuclear extract，核提取物
NELF，negative elongation factor，负延伸因子
NF-κB，nuclear factor-κB，核因子-κB
NFAT，nuclear factor of activated T-cells，活化 T 细胞核因子
NHPs，nonhistone proteins，非组蛋白

Ni-NTA，nickel-nitrilotriacetic acid，镍-次氮基三乙酸
NMR，nuclear magnetic resonance，核磁共振
NP-40，Nonidet P-40，诺乃洗涤剂 P-40
NPE，nucleosome positioning sequence，核小体定位序列
NR，nuclear receptor，核受体

OD，optical density，光密度
ONPG，*O*-nitrophenyl-β-D-galactopyranoside，邻硝基苯-β-D-半乳糖苷
OP-Cu，*O*-phenanthroline-$CuSO_4$，*O*-邻菲啰啉-硫酸铜
ORF，open reading frame，可读框

PADI4，peptidylarginine deiminase 4，肽基精氨酸脱亚氨酶 4
PAGE，polyacrylamide gel electrophoresis，聚丙烯酰胺凝胶电泳
PAH2，paired amphipathic helix 2 domain，配对两性螺旋 2 结构域
PBS，phosphate-buffered saline，磷酸缓冲盐
PBV，packed bead volume，装珠量
Pc，polycomb，多梳
PCR，polymerase chain reaction，聚合酶链反应
PEAS，*N*-［（2-pyridithio）ethyl］-4-azidosalicylamide，*N*-［（2-吡啶基硫代)乙基］-4-叠氮水杨酰胺
PEG，polyethylene glycol，聚乙二醇
PEI，polyethyleneimine，聚乙烯亚胺
PHD，plant homeodomain，植物同源域
PlC，preinitiation complex，起始前复合物
PKA，protein kinase A，蛋白激酶 A
PML-RARcL，promyelocytic leukemia-retinoic acid receptor-C，早幼粒细胞白血病维甲酸受体-C
PMSF，phenylmethyl sulfonyl fluoride，苯甲基磺酰氟
Pol II，RNA polymerase II，also abbreviated RNAP II，RNA 聚合酶 II，也简称 RNAP II
PRC1，polycomb repressive complex 1，多梳抑制复合物 1
PRC2，polycomb repressive complex 2，多梳抑制复合物 2
PREs，polycomb repressive elements，多梳抑制元件
PRMT，protein arginine（R）methyltransferase，蛋白质精氨酸（R）甲基转移酶
PSA，prostate-specific antigen，前列腺特异性抗原
PVA，polyvinyl alcohol，聚乙烯醇

RACE，rapid amplification of cDNA end，快速扩增 cDNA 末端

TEMED，N，N，N′，N′-tetramethylethylenediamine，N，N，N′，N′-四甲基乙二胺
TEV，tobacco etch virus，烟草蚀纹病毒
TFII，transcription factor II，转录因子 II
Th2，T-helper 2，T 辅助 2
TLC，thin layer chromatography，薄层色谱法
T_m，melting temperature，熔解温度
Topo I，topoisomerase I，拓扑异构酶 I
TR，thyroid hormone receptor，甲状腺激素受体
TRAP，thyroid hormone receptor-associated protein，甲状腺激素受体相关蛋白
Trp-R，tryptophan operon repressor，色氨酸操纵子阻遏
TSS，transcription start site，转录起始位点

Ub，ubiquitin，泛素蛋白
USA，upstream stimulatory activity，上游刺激活性
USF，upstream stimulatory factor，上游刺激因子

VDR，vitamin D receptor，维生素 D 受体
vdw-h，van der Waals contacts with a C5 hydrogen on cytosine，胞嘧啶 C5 氢范德华接触
vdw-me，van der Waals contacts with the methyl group projected from the C5 of thymine，胸腺嘧啶 C5 甲基范德华接触
VSV，vesicular stomatitis virus，水泡性口炎病毒

XPB，xeroderma pigmentosum B，着色性干皮病 B

ZEBRA，Z Epstein-Barr replication activator，Z 爱泼斯坦-巴尔复制激活性因子

目　录

1　哺乳动物细胞转录调控入门

引言和概述

蛋白质组、分子、细胞和全基因组研究的结合，揭示了很多关于真核生物转录差异调控机制的认识。然而，这一领域才仅仅开始瞥见调控的复杂性，即使在酵母一样简单的生物中也是如此。在人类中，庞大的转录因子网络控制着转录波，而这种转录波导致发育、分化及信号转导期间约 25 000 个基因的差异转录。这种转录涉及 RNA 聚合酶 II (Pol II)、被称为调解因子的共激活因子，以及 6 种被称为 TFIIA、TFIIB、TFIID、TFIIE、TFIIF、TFIIH（见下面）的辅助因子的募集。这些因子通过一个明确的循环过程使转录起始和再起始（图 1.1）。

正如荧光原位杂交（FISH）所测定的，染色体似乎在细胞核内占据不同的区域（染色体域，CT），并且已经猜测染色体间通讯发生的原因是大的染色质环的动态移动（Kadauke and Blobel 2008），这种移动可能由肌动蛋白或核肌凝蛋白 I 所控制（Branco and Pombo 2006；Cremer et al. 2006）。据认为，转录在 CT 内和 CT 间存在的 200～

2000 个转录工厂的任何一个中发生。

在所有的真核生物中，染色体 DNA 包装成染色质。染色质分为异染色质和常染色质。异染色质复制晚、着色深、致密，并经常出现在细胞核边缘、靠近核仁或染色质中心。通常认为异染色质是转录沉默的，然而最近该领域中的研究表明异染色质中发生有限的转录。常染色质染色浅、松散，并且通常但也不绝对地定位在细胞核内部。常染色质复制早，且含有基因组转录活跃区域。兼性染色质是异染色质，它有转变成常染色质的能力，反之亦然（Trojer and Reinberg 2007）。

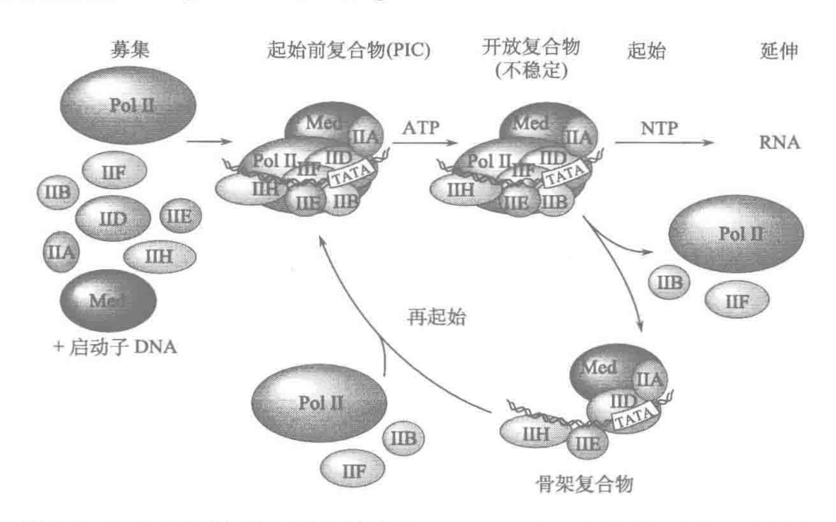


图 1.1　Pol II 转录起始和再起始途径（经 Macmillan 出版公司允许，改编自 Hahn 2004）。

染色质的基本单位是核小体，核小体是由 147 bp DNA、各 2 个拷贝的 4 种核心组蛋白，以及在某些环境中的一个连接组蛋白组成的（Horn and Peterson 2002；Luger Hansen 2005；Woodcock 2006）。很多组蛋白有数个变体和异构体，它们似乎在基因调控中发挥特定作用。核小体自身安排成高级结构并受多种形式的调控。基因活化和沉默过程是由序列特异性 DNA 结合蛋白精细编排的，这些蛋白质募集涉及改变染色质模板和转录基因的酶及结构蛋白（structural protein）。估计在人类基因组中，编码 1800～1900 个序列特异性结合蛋白，远不及基因个数或 mRNA 表达模式。因此，即使人类最简单的基因也是由数个序列特异性因子组合调控的（Maston et al. 2006）。激活因子以协作方式刺激转录，以至于基因对单一调控信号不应答，但对控制不同激活因子的数种聚集信号作出应答。这种组合式调控类型可能是基因差异表达中很重要的方面之一（Carey 1998）（见第 12 章）。

本章是对转录和染色质调控机制的基本概述。具体论题在随后的章节中进一步讨论。若有可能，我们选择引用综述性文章，因为涵盖的论题涉及大量原始文献，并且该领域中的新手将可能需要对本章所讨论的每个方面的一般性介绍。本章仅仅聚焦于聚合酶 II（Pol II）的转录。已出现了关于 Pol I 和 Pol III 的优秀综述，Pol I 转录编码大核糖体 RNA 的 45S 前体基因，而 Pol III 转录很多小 RNA 基因，包括为 tRNA 和 5S 核

糖体（Haeusler and Engelke 2006；Russell and Zomerdijk 2006；Chedin et al. 2007；Dieci et al. 2007；Cramer et al. 2008）。在植物中，第 4 种 RNA 聚合酶在合成微 RNA（microRNA，miRNA）时出现（Vaughn and Martienssen 2005），但是像线粒体这样的细胞器有它们自己的 RNA 聚合酶（Scarpulla 2008）。

全基因组方法小结

方法的简要概述可能有助于理解在本入门章节中将出现的一些信息，尽管第 9 章中我们也将涵盖这些方法中的部分方法。通常通过染色质免疫沉淀来研究全基因组蛋白质-DNA 相互作用（Kim and Ren 2006），其中，染色质免疫沉淀使用来自甲醛处理的细胞核的片段化染色质。甲醛将蛋白质交联到 DNA 上，并且通过超声或核酸酶处理使交联的蛋白质-DNA 复合物片段化，然后用靶向目的蛋白或组蛋白修饰的抗体对蛋白质-DNA 复合物进行免疫沉淀。之后，分离沉淀的 DNA、解交联，并使用荧光标记的核苷酸对 DNA 进行聚合酶链反应（PCR）扩增。然后将扩增的 DNA 与带有叠连 DNA 阵列（tiled DNA array）的芯片杂交，这些叠连 DNA 阵列代表了启动子（启动子阵列）及基因组感兴趣的区域，包括某些启动子和增强子［如 ENCODE 阵列，ENCODE 计划联盟（ENCODE Project Consortium）2004］；或者，可构建用户定制设计的阵列。通过芯片荧光信号测定的杂交量是蛋白质与 DNA 结合的一种度量，通过荧光阅读器对这种杂交荧光信号进行检测并使用附带的软件对其定量。这些所谓的 ChIP-chip（或 ChIP-on-chip）分析已在该领域广泛用于破译控制转录的调控线路（图 1.2）。

ChIP-chip 技术有很多局限性，包括抗体、覆盖整个基因组的费用，以及在选择要检查的区域中产生的偏向性。另外，ChIP 不是定量的，难以把占据率的测量结果归因于 ChIP。例如，对任何特定的基因，如果具有该基因的所有细胞或仅一小部分细胞包含结合的蛋白质或组蛋白修饰，那么当前用 ChIP 来确定是不可行的。在两种修饰或结合事件出现在相同位置的情况下，有可能对一个样品再次 ChIP，以确定它是否含有另一种修饰，这种操作很少在全基因组研究中进行。此外，ChIP 主要是描述性的，并不产生机制信息。尽管如此，ChIP 产生的信息可能具有启迪作用，并使研究者提出可测试的假设。

最新的技术涉及 ChIP 及随后的微流体装置中沉淀 DNA 的高通量 DNA 测序（ChIP-Seq 或 ChIP Sequencing）（Collas and Dahl 2008）。该方法具有无偏的优点，但受限于抗体质量和全基因组覆盖；虽然全基因组覆盖仍然是困难的，但已经可以实现，该技术及增加序列覆盖度的能力正在迅速提高。这些技术使一些简单分析，如转录因子和组蛋白修饰的 ChIP 分析，能够在全基因组基础上进行。

ChIP-chip 和 ChIP-Seq 技术可适用于 DNA 甲基化研究（Zilberman and Henikoff 2007）、DNase I 高敏感性（Crawford et al. 2006a，b；Xi et al. 2007）、染色质构象捕获（3C）（Simonis et al. 2007）及 RNA 陷阱（Chakalova et al. 2004）。

可通过以甲基化 CpG 抗体沉淀染色质来测量 DNA 甲基化。DNase I 高敏感性是测量染色质对 DNase I 可及性（accessibility）的一种技术。以增强子、启动子和绝缘子为特征的开放染色质对切割更敏感。对于单个基因，通过 Southern 印迹分析 DNase I 切割的 DNA，或者将 DNase I 切割的 DNA 连接接头引物，使用接头引物扩增 DNA，用

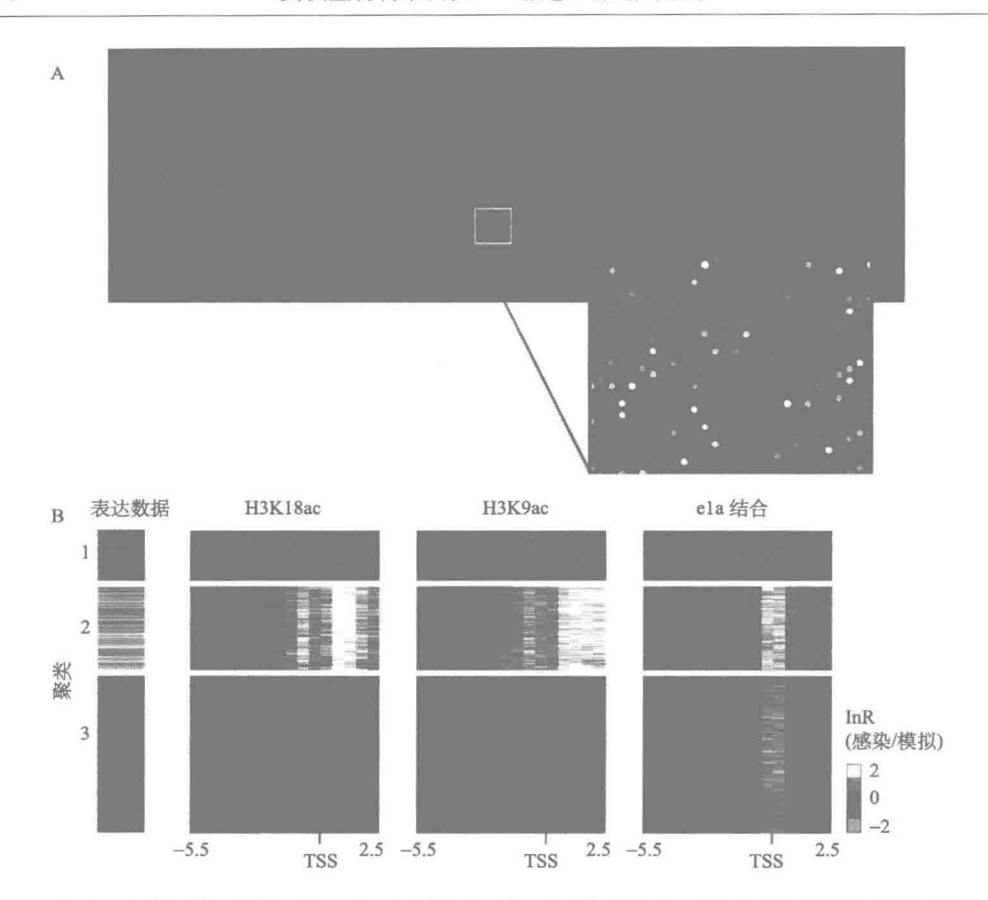


图 1.2　全基因组 ChIP-chip。在一般的实验中，通过 PCR 对来自 ChIP 反应的 DNA 进行扩增，扩增中以荧光团标记扩增 DNA，扩增产物同微阵列杂交。A. 安捷伦（Agilent）启动子微阵列的扫描图像，该微阵列同来自腺病毒感染细胞的荧光标记 ChIP DNA 杂交。插入图显示了该芯片的一个小区域的放大图像。这些阵列由 2 张片子组成，共包含 17 054个注释的人类启动子的探针，覆盖了每个启动子相对于已知注释的转录起始位点（TSS）−5.5～+2.5 kb 的区域。每片微阵列包含 244 000 个 60 bp 长的寡核苷酸探针。每个点的荧光强度反映 ChIP 样品中相应 DNA 片段的富集水平。将从芯片扫描获得的数据归一化，并通过包括聚类算法的各种统计方法对其进行分析（见下文）。该插图是灰度图，但一般以绿色和红色显示阵列图像。B. 该图显示了灰度热图，但通常用绿和红或者黄和蓝显示热图。根据给定的比例，热图用于直观显示每个基因座位的富集水平。按照惯例，红或黄色调增强反映更高富集值，而绿或蓝色调的增强反映相对耗尽。在显示的该例子中，按照 K-均值排列数据，可见这些数据聚类为三组，聚类的依据是基因表达数据和基因个体发生学分类（http://david.abcc.ncifcrf.gov/）。沿垂直轴的行各代表一个单独的基因，并且根据右下角的比例尺，各列反映 ChIP 信号。每列代表沿启动子长度的 500 bp 间隔，启动子区从−5.5 到+2.5 kb。在该实验中，用腺病毒感染细胞，在感染 24 h 后，通过基因表达微阵列检测 mRNA 水平，并且通过 ChIP-chip 检测 H3K18 和 H3K9 的乙酰化水平，以及腺病毒小 E1A（e1a）蛋白的结合。图版从左到右依次表示基因表达数据（Exp）、腺病毒感染时 H3K18 和 H3K9 的乙酰化水平，以及 E1A 结合。注意：E1A 的结合主要发生在 TSS 周围，但它对组蛋白乙酰化的影响从 TSS 上游和下游扩展到基因的转录区（A 由 R. Ferrari 提供；B 经 AAAS 许可，改编自 Ferrari et al. 2008）。

于测序或 ChIP 分析。

3C 是将甲醛交联过的蛋白质-DNA 复合物用限制酶消化并稀释，然后加入 DNA 连接酶的一种技术。如果一个蛋白质交联到相距较远的两段 DNA（即一个增强子和启动子）上，那么这两个 DNA 片段将被连接在一起，因为它们在空间上相互接近（即交联到相同的蛋白质上）。之后，用特异性引物或者连接接头扩增 DNA 片段。所得到的 DNA 片段可与芯片杂交或测序。

RNA 陷阱方法目前仅有少数实验室使用，该技术设计互补于一个 RNA 分子的寡核苷酸探针，其中，RNA 分子代表一个接近转录起始位点（TSS）的区域。该探针同隔离的染色质上的 RNA 杂交，其中该染色质上包含正在延伸的 Pol II。将寡核苷酸探针连接到地高辛上，地高辛与辣根过氧化物酶偶联的地高辛抗体结合。加入生物素酪胺，并且过氧化物酶产生自由基，自由基将生物素连接到附近的蛋白质上。DNA 片段化后，在链亲和素偶联的树脂上分离生物素标记的蛋白质和 DNA。捕获的 DNA 片段可以扩增并测序，或同芯片杂交。因此，如果相距较远的两段 DNA 接近该 RNA 的 5′端，它们将在该程序中被分离。

下面要讨论的另一种技术是 FISH，这种方法通常在固定的细胞核或细胞上进行，在该方法中，荧光标记的核苷酸探针同 RNA 或 DNA 杂交，并且使用共聚焦显微镜监测荧光。将 FISH 同免疫荧光结合起来，比较 DNA 和 RNA 相对于各种含蛋白质结构的位置。该技术用于定位染色体域，还可鉴定 mRNA 相对于含 RNA Pol II 的核点（foci）的位置，其中含 RNA Pol II 的核点被称为转录工厂（transcription factory）。此外，在某些情况下，该技术还可用于研究诸如基因座控制区（LCR）的 DNA 元件相对于正在远处活跃转录的基因的位置。

染色质和通用转录机器

染色质结构和组织

染色质凝聚使得 2 m 长的 DNA（直径 2 nm，每圈 10.4 bp、高 3.4 nm）适合 5～20 μm 直径的细胞核。细胞核的 FISH 分析已经揭示出在一般的哺乳动物间期细胞核中，每个染色体占据不同的区域，该区域称为 CT（Branco and Pombo 2006；Cremer et al. 2006）。这些 CT 内的染色体由基本的 10 nm 直径的核小体纤维（串珠结构）组装而成，其中，每个核小体包含一个核心组蛋白八聚体（包含各两分子的组蛋白 H2A、H2B、H3 和 H4），以及偶尔出现的一个连接组蛋白——H1 或 H5（Horn and Peterson 2002；Luger and Hansen 2005；Woodcock 2006）。10 nm 核小体纤维再组装成称为 30 nm纤维的高级结构（Tremethick 2007；Wu et al. 2007），30 nm 纤维卷曲成 80～100 nm 的染色丝纤维（Horn and Peterson 2002）。有可能存在数种其他层次的组织，包括染色体域内或染色体域间的多种形式的染色质环、异染色质区，以及转录工厂（Kadauke and Blobel 2008）。除了组蛋白，包括各种高迁移组（HMGA、HMGB 和 HMGC）蛋白（Hock et al. 2007）的几种高丰度非组蛋白（NHP）通常与染色质结合，并与核基质附着位点一起用于染色质的组织（Linnemann et al. 2007）。核基质本

身是一组不溶性蛋白，构成附着于核纤层的纤维网格。

最初的高分辨率核小体晶体结构是在 1997 年解析的，此后，出现了更高分辨率的核小体结构，除此之外，还出现了含有数个组蛋白变种的核小体结构（见下文）(Chakravarthy et al. 2005)。此外，已解析了四联核小体结构，从而洞悉了高级核小体包装可能是如何发生的（Schalch et al. 2005；Tremethick 2007）。一个典型的核小体包含 146～147 bp 的 DNA，该 DNA 以环形超螺旋缠绕在组蛋白八聚体周围，共缠绕 1.65 次。该结构揭示核小体具有约 6 nm×10 nm 的类圆盘形状。可将组蛋白八聚体细分为一个 H3/H4 异源四聚体和两个 H2A/H2B 异源二聚体（图 1.3）。

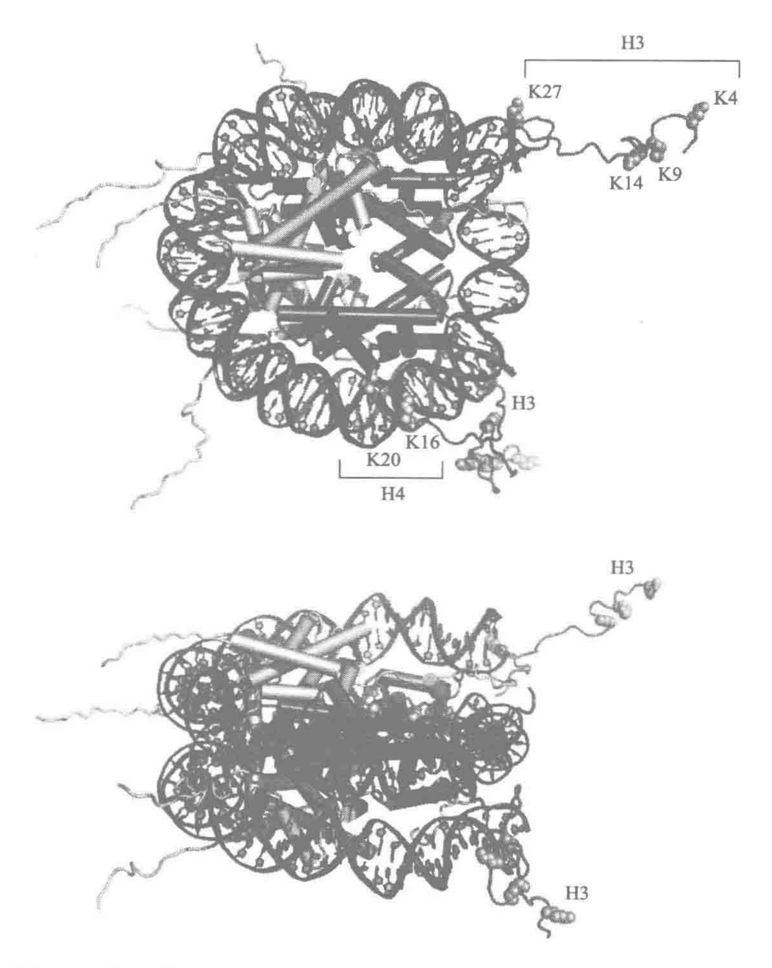


图 1.3　核小体的两种视图。核小体用 PyMOL 和 PDB file 1KX5 绘制。尾巴是人为从核小体核心延伸出来的。(深灰色) H3/H4 四聚体；(浅灰色) H2A/H2B 二聚体。下述尾巴残基突出显示：H3 上的 Lys 4、9、14 和 27，以及 H4 上的 Lys 16 和 20。注意，H4 尾巴比 H3 短。因此，残基 16 和 20 非常靠近起始的 H4 α 螺旋且靠近 DNA（由 J. Heiss 绘制，UCLA）。

每个组蛋白含有一个特征性模体，称为组蛋白折叠。该模体由一个两端通过短的β转角与两个短α螺旋连接的长α螺旋组成。组蛋白将精氨酸插入到所有14个面向单核小体中八聚体表面的DNA小沟中。特异性接触点的个数估计约为120个。这些接触提供了驱使DNA-八聚体相互作用的大部分能量。依据单个二核苷酸碱基阶梯的可变形性，不同的DNA序列有不同的折叠进核小体的倾向（Richmond 2006；Segal et al. 2006)。自相矛盾的是，生物物理学的研究表明，核小体中DNA-蛋白质的相互作用是动态的。在生理性盐浓度下，核小体中DNA的出入点似乎瞬时脱离八聚体的表面，甚至在诸如人工601序列这样的高亲和力核小体定位序列上也是如此（Mihardja et al. 2006；Hoch et al. 2007)。此外，在低亲和力序列上，核小体可沿DNA滑动，诸如NAP1的组蛋白分子伴侣（Park and Luger 2006）和三磷酸腺苷（ATP）依赖性重塑机器可促进这种滑动（Cairns 2005；Saha et al. 2006b)。连接组蛋白H1和H5在氨基末端包含一个两侧有非结构化结构域的翼状螺旋结构域，并且在羧基末端含有高碱性但非结构化的区域。氨基末端被磷酸化，并且氨基和羧基末端及翼状螺旋结构域在数个位置被翻译后修饰（Woodcock et al. 2006；Wisniewski et al. 2007)。据认为，连接组蛋白在核小体的出入点结合DNA，但这是否总是正确的，还有待于进一步的研究。连接组蛋白的结合限制核小体的移动性。

所有4种组蛋白都包含富含赖氨酸的氨基末端尾巴。此外，H2A组蛋白有类似的富含赖氨酸和丝氨酸的羧基末端尾巴，该尾巴未被很好地表征。组蛋白尾巴从核小体的表面伸出，并且在组织或稳定核小体DNA中不发挥主要作用。然而，这些尾巴经受很多翻译后修饰，这些修饰在转录、复制和修复期间充当控制染色质代谢的信号通路的组件。某些尾巴稳定核小体间的DNA相互作用。例如，H4的尾巴同相邻核小体内H2A/H2B二聚体的暴露表面相互作用，该表面主要是由来自H2A的7个酸性残基形成的。有证据表明，H4的尾巴对30 nm纤维的形成是必要的，因为H4尾巴在K16的乙酰化导致赖氨酸碱性电荷的中和，这解开了H4尾巴与H2A/H2B二聚体的接触并阻止30 nm纤维的形成（Shogren-Knaak et al. 2006)。

有很多从组蛋白主要家族中分出的组蛋白变种（Tin et al. 2005；Bernstein and Hake 2006；Hake and Allis 2006；Workman 2006；Eirin-Lopez and Ausio 2007；Kusch and Workman 2007；Loyola and Almouzni 2007)，这些变种包括H2AZ、H2AX、macroH2A、H2A. Bbd、H3. 3及CenH3。发现H2AZ以未乙酰化的形式存在于基因组的抑制区，而当H2AZ被乙酰化时，则存在于基因组的活性区（即启动子)。磷酸化的H2AX出现在DNA损伤位点。macroH2A富集于雌性无活性X染色体中。H2A. Bbd位于染色体的转录活性区，并且被排除在无活性区（即排除于无活性的X；Bbd代表巴氏体缺失)。H3. 3在活跃转录期间被整合进基因中，而CenH3或CenpA位于着丝粒异染色质内。

关于这些替代组蛋白的性质仍然有很多要了解，但它们的功能可能与它们的定位有关。例如，通过SWR1 ATP重塑酶，组蛋白变种H2AZ（约60%同H2A一致）被插入酿酒酵母（*S. cerevisiae*）和哺乳动物启动子两侧的核小体中（Korber and Horz 2004)。含有H2AZ的核小体包含增加30nm纤维形成的被延伸的酸性补丁（McBryant

et al. 2006；Tremethick 2007)，但是含有 H2AZ 的核小体可能不及含有 H2A 的核小体稳定（Zhang et al. 2005）。尽管如此，结构分析揭示 H2AZ 被纳入核小体中，而没有任何明显的 DNA 接触的扭曲或损失（Chakravarthy et al. 2005）。

30 nm 细丝的结构是目前许多研究的主题，因为相关的证据有力地表明细丝内的基因必须被“解压缩”才能胜任转录。在 30 nm 细丝的早期电子显微镜（EM）研究中，认为单个核小体呈螺线管构型，折叠成所谓的单头螺旋。该结构的证据包括 EM 和交联数据。相比之下，最近的高多分辨率数据表明其还有一种替代形式，称为 Z 形双头螺旋（图 1.4）（Tremethick 2007；Wu et al. 2007）。该模式中，两个核小体彼此间弯曲成 Z 形，与交叉在二者间的连接 DNA 形成两行。例如，第一个和第三个核小体被堆叠在一行，并且通过连接 DNA 连接到第二行上，而第二行上堆叠了第二个和第四个核小体。在高镁离子浓度中形成的四联核小体的晶体结构揭示了核小体的两个堆叠，同时在两个堆叠之间带有 Z 形弯曲的连接 DNA，从而表明与 30 nm 纤维的一种关系，即使该模型显示直径可能不是 30 nm。通过 EM 研究显示，含有 H1 的核小体形成类似的 Z 形结构。连接组蛋白的整合可显著地稳定 30 nm 细丝。最后，短核小体阵列的交联数据也支持 Z 形结构。

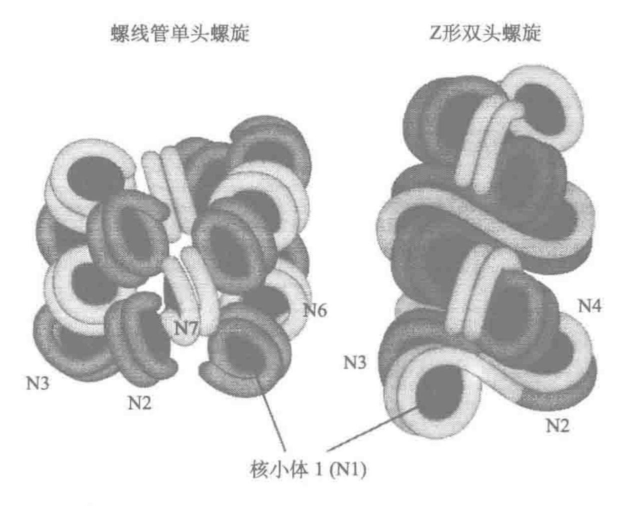


图 1.4　30 nm 细丝的结构模式。显示 30 nm 细丝的两种模式：螺线管单头螺旋和 Z 形双头螺旋。当前有支持两种模式的数据（经 Elsevier 许可，改编自 Trementhich 2007）。

染色质修饰

组蛋白尾巴的修饰是转录活化和沉默期间调控的主要形式（Kouzarides 2007），该尾巴及某些情况下组蛋白的球状结构域，受到至少 8 种类型的共价标记：乙酰化、甲基化（赖氨酸单甲基化、双甲基化和三甲基化，以及精氨酸单甲基化和双甲基化，其中双甲基化可对称和不对称）、磷酸化、泛素化、SUMO 化、二磷酸腺苷（ADP）-核糖基化、脱亚氨基作用及脯氨酸异构化。这些标记具有环境依赖性的效应，同时，在生物体

内存在除去修饰的酶，包括组蛋白去乙酰化酶（HDAC）、组蛋白去甲基化酶和去泛素化酶。因此，修饰是动态的，并且当特定的过程必须逆转时（即基因激活和抑制/沉默期间），要去除修饰。催化修饰的酶在不同的调控环境中起作用，并参与染色质代谢的所有方面。通过质谱和ChIP技术，已经检测到很多修饰，其中有些修饰呈复杂的组合。

表 1.1 染色质的共价修饰

组蛋白修饰酶	被修饰的残基	组蛋白修饰酶	被修饰的残基
乙酰转移酶		SUV4 20H2	H4K20
HAT1	H4（K5、K12）	SpSet 9	H4K20
CBP/P300	H3（K14、K18）H4（K5、K8）	EZH2	H3K27
	H2A（K5）H2B（K12、K15）	RIZ1	H3K9
PCAF/GCN5	H3（K9、K14、K18）		
TIP60	H4（K5、K8、K12、K16）	赖氨酸去甲基转移酶	
	H3 K14	LSD1/BHC110	H3K4
HB01（ScESA1、SpMST1）	H4(K5、K8、K12)	JHDM1a	H3K36
ScSAS3	H3(K14、K23)	JHDM1b	H3K36
ScSAS2（SpMST2）	H4 K16	JHDM2a	H3K9
ScRTT109	H3 K56	JHDM2b	H3K9
		JMJD2A/JHDM3A	H3K9、H3K36
赖氨酸甲基转移酶		JMJD2B	H3K9
SUV39H1	H3K9	JMJD2C/GASC1	H3K9、H3K36
SUV39H2	H3K9	JMJD2D	H3K9
G9a	H3K9	UTX	H3K27
ESET/SETDB1	H3K9	JMJD3	H3K27
EuHMTase/GLP	H3K9		
CLL8	H3K9	精氨酸甲基转移酶	
SpClr4	H3K9	CARM1	H3（R2、R17、R26）
MLL1	H3K4	PRMT4	H4R3
MLL2	H3K4	PRMT5	H3R8、H4R3
MLL3	H3K4		
MLL4	H3K4	丝氨酸/苏氨酸激酶	
MLL5	H3K4	Haspin	H3T3
SET1A	H3K4	MSK1	H3S28、H3S10
SET1B	H3K4	MSK2	H3S28、H3S10
ASH1	H3K4	RSK2	H3S10
Sc/Sp SET1	H3K4	CKII	H4S1
SET2（Sc/Sp SET2）	H3K36	Mst1	H2BS14
NSD1	H3K36		
SYMD2	H3K36	泛素化酶	
DOT1	H3K79	Bmi1/Ring1A/Ring2B	H2AK119
Sc/Sp DOT1	H3K79	RNF20/RNF40	H2BK120
Pr-SET 7/8	H4K20		
SUV4 20H1	H4K20	脯氨酸异构酶	
		ScFPR4	H3P30、H3P38

经Elsevier许可，改编自Kouzarides（2007）。

表中已包括了只对一个或几个位点具有特异性的酶，同时包括了它们修饰的位点。表中显示了人和酵母中的酶。酵母酶加前缀Sc（酿酒酵母，*Saccharomyces cerevisiae*）或Sp（裂殖酵母，*Schizosaccharomyces pombe*）以示区分。按属于相同家族的酶进行分组。注意，新的系统命名法已被认可用于命名染色质修饰酶（Allis et al. 2007）。

染色质修饰有两个整体性的作用。一个作用是为一些蛋白质产生结合位点，这些蛋白质含有用于特定修饰的识别结构域，然后，募集到该修饰上的蛋白质引发一种效应，下文会讨论这些效应的几个例子。另一个作用是，如上文对 H4K16 乙酰化所做的讨论，某些修饰会影响染色质压缩。组蛋白中的某些突变体，如 Sin（不依赖 SWI/SNF 的）突变体也会影响核小体的稳定性和染色质压缩。表 1.1 列出了负责各种修饰的酶，而表 1.2 列出了参与结合这些修饰的结构域的类型，下面的几节中将讨论具有代表性的这些酶类及其结构域。

表 1.2　修饰识别结构域

读取器模块	PTM 标记
皇家家族	
溴域	很多组蛋白 Kac，（Kac）
染色域	H3K9me2/3、H3K27me2/3
双染色域	H3K4mel/2/3
染色桶结构域	H3K36me2/3
Tudor	（Rme2s）
双/串联 Tudor	H3K4me3、H4K20me3、H4K20mel/2，（Kme2）
MBT	H4K20mel/2、HlK26mel/2、H3K4mel、H3K9mel/2
PHD 指	H3K4me3、H3K4me0、H3K9me3、H3K36me3
WD40 重复	H3R2/K4me2，（R、Sph、Tph）
14-3-3	H3S10ph、H3S28ph，（Sph、Tph）
BRCT	H2AX-S139ph，（Sph、Tph）

经 Macmillan Publishers Ltd. 许可，改编自 Taverna 等（2007）。

按结构域家族分组显示已知的染色质结合模块，以及已报道的这些模块结合的组蛋白标记。括号表示例子，其中，相关家族成员及其与非组蛋白 PTM（翻译后修饰）的相互作用的结构信息是已知的。

乙酰化

乙酰化是用乙酰辅酶 A(CoA) 作为供体，通过组蛋白乙酰转移酶（HAT）将乙酰基转移到赖氨酸上实现的（Dyda et al. 2000；Peterson and Laniel 2004；Lee and Workman 2007）。在该反应中，带正电荷的 NH_3^+ 被不带电荷的 $NH—CO—CH_3$ 取代。对表 1.1 的粗略检查揭示出关于各种修饰酶特异性的显著观察结果。很多 HAT 作用底物混乱并靶向很多赖氨酸残基的乙酰化，有时在不同组蛋白中乙酰化。例如，正如在第 10 章中的进一步讨论，在哺乳动物细胞中，高度相关的 p300 同 CREB 结合蛋白（CBP，通常称为 p300/CBP），修饰所有 4 种组蛋白尾巴上的残基。相关的哺乳动物 GCN5 和 PCAF 酶，以及酿酒酵母的 Gcn5（在 SAGA 中发现）修饰 H3 上数个残基。值得注意的是，正如通过单个研究和全基因组分析所测定的，在活跃启动子中发现数个残基在哺乳动物细胞中被 p300/CBP 和 GCN5/PCAF 修饰，下文还要对此进行讨论（Wang et al. 2008）。

一种已知结构的专门结构域识别乙酰化标记，这种结构域被称为溴域（Mujtaba et

al. 2007)。在大量HAT和ATP依赖性重塑复合物的亚基中发现了溴域。相反，有许多含有溴域的因子并不是ATP依赖性机器的亚基，如酿酒酵母中的溴域因子1（BDF1），据认为，BDF1相当于在人（而不是酿酒酵母）TATA盒结合蛋白相关因子1（TAF1；见下文）中发现的一个缺失片段。另外，在哺乳动物细胞中的溴域蛋白2（BRD2）和BRD3与H4K5ac、H4K12ac及H3K14ac结合，并且可能通过担当组蛋白伴侣分子，在体外通过染色质刺激Pol II的转录（LeRoy et al. 2008）。

甲基化

组蛋白甲基化发生在赖氨酸和精氨酸上（Cheng et al. 2005；Martin and Zhang 2005，2007；Shilatifard 2006）。在赖氨酸上，甲基化转移酶（甲基化酶）催化*S*-腺苷甲硫氨酸上的甲基转移到赖氨酸上，取代NH_3^+上的1、2或3个氢（单、双和三甲基化）。同乙酰化酶相比，大多数组蛋白甲基化酶和去除甲基的酶，即去甲基化酶，具有强烈的特异性。哺乳动物中6种不同的赖氨酸被单、双及三甲基化，包括H3K4、H3K9、H3K27、H3K36、H3K79及H4K20。

在哺乳动物细胞中，在臂间异染色质形成和常染色质基因沉默的过程中，SUV39H1和SUV39H2三甲基化H3K9，而在常染色质基因沉默过程中，G9a二甲基化H3K9。涉及沉默的其他甲基化酶包括SUV4-20H1和EZH2酶；SUV4-20H1酶催化异染色质中H4K20的三甲基化，而EZH2酶催化兼性异染色质（即在失活的*Hox*基因上）中H3K27的三甲基化。

在转录起始和延伸的早期阶段，MLL1～5家族成员及SET1A和SET1B高度特异地催化H3K4双甲基化和三甲基化（Eissenberg and Shilatifard 2006；Shilatifard 2008）。类似地，在基因活化及延伸期间，与Pol II结合的SET2使H3K36甲基化，并且DOT1修饰H3K79。H3K36me3似乎局限在H3K4me3的下游，而H3K79me3遍及整个基因。除了DOT1，其他甲基化酶是甲基化酶的SET结构域家族的一部分，并且共享结构上保守的催化结构域（Li et al. 2007a）。蛋白质精氨酸（R）的甲基化酶（PRMT）使精氨酸甲基化。这些酶修饰H3R2、R8、R17、R26及H4R3，并且有可能修饰其他位置（Cheng et al. 2005；Kouzarides 2007）。

有数种类型的结构域识别甲基化的组蛋白尾巴（Taverna et al. 2007），这些结构域包括皇家家族的成员，如染色域、双染色域、染色桶结构域、Tudor结构域、双Tudor结构域和恶性脑瘤（MBT）结构域。植物的同源结构域（PHD）和WD40重复也能结合特定的甲基化赖氨酸或精氨酸。带有这些结构域的蛋白质靶向各种修饰。例如，HP1的染色域指导蛋白质结合H3K9me2和H3K9me3，而多梳蛋白（Pc或人HPC）的染色域结合H3K27me3。这些例子及其他例子在随后几节中讨论。

去甲基化是由称为jumonji家族和胺氧化酶家族的两类去甲基化酶完成的，其中，胺氧化酶家族以赖氨酸特异性去甲基化酶1（LSD1）为代表（Anand and Marmorstein 2007；Shi and Whetstine 2007；Lan et al. 2008）。去甲基化酶可分为单、双或三甲基特异性去甲基化酶，但由于LSD1的化学断裂机制，该酶只能进行双甲基化和单甲基化组氨酸的去甲基化。去甲基化酶比甲基化酶的作用对象可能更加复杂，并且有时能够将两个不同的残基去甲基。例如，jumonji去甲基化酶JMJD2A-D可以使H3K9me2/me3

和 H3K36me2/me3 去甲基。

该领域的关键问题是这些及其他组蛋白修饰酶是如何被募集到基因上的。UTX 和 JMJD3 是 H3K27me2 和 H3K27me3 的去甲基化酶。这些蛋白质是含有 MLL2 和 MLL3 H3K4 甲基化酶的蛋白复合物的组分（Lan et al. 2008）。因此，带有与基因活化相关的甲基化酶的复合物包含去除抑制标记的酶。相似地，下文描述的 CoREST 和 CtBP 共抑制因子复合物含有与 HDAC 一起作为亚基的 LSD1。在这种情况下，由序列特异性抑制因子募集的复合物至少包含两种酶，这两种酶能够除去与转录活化相关的修饰。

单泛素化

发生在组蛋白赖氨酸上的单泛素化（Ub）是参与基因调控的另一种修饰方式（Shilatifard 2006），它与基因活化及基因沉默均有关系。例如，多梳抑制复合物 1（PRC1）的 Ring1B/Ring1A/Bmi 结合亚基使 H2AK119 泛素化，PRC1 带有含染色域的 Pc 或 HPC 蛋白（见下文）。在这种情况下，一个标记，即 H3K27me3，可募集一种复合物，该复合物加入第二种参与沉默的抑制标记。相比之下，H2BK120Ub（或酿酒酵母中 H2B123Ub）在人类中被 RNF20/40 和 Ubc6 蛋白催化，而在酿酒酵母中被 Rad6/Brel 复合物催化。正如下文所描述的，在酿酒酵母中的延伸起始阶段，不知何故，该修饰对 H3K4 的三甲基化是必需的。在哺乳动物染色质转录系统中，H2B120Ub 强烈刺激体外 Pol II 延伸（Pavri et al. 2006）。自相矛盾的是，在酿酒酵母中，带有 Gcn5 HAT 的 SAGA 复合物包含一个称为 Ubp8 的亚基，该亚基使 H2BK123 去泛素化。因为泛素化和去泛素化均与基因活化有关，这可能表明一系列相继事件对激活是必需的（Shilatifard 2006）。

磷酸化

组蛋白磷酸化也有助于基因活化（Kouzarides 2007）。H3S10 的磷酸化是与有丝分裂和哺乳动物中很多快速应答基因的活化有关的一种标记。哺乳动物中 MSK1/2 和 RSK2 激酶使 H3S10 磷酸化，在酿酒酵母中 Snf1 使 H3S10 磷酸化。在哺乳动物细胞中的 DNA 损伤期间，H2AX S139 被 ATM 激酶磷酸化并产生所谓的 γH2AX。各种 14-3-3 蛋白识别 H3S10P，并且 DNA 损伤检测点蛋白 1（MDC1）的调解因子的串联 BRCT1 结构域识别 γH2AX（磷酸化的）。在另一个例子中，在有丝分裂期间，Aurora B 激酶使 H3S10 和 H3S28 磷酸化。除了募集其他蛋白质外，H3S10P 还可影响自身。据认为，H3S10P 尤为重要，因为它可以阻止 HP1 蛋白与 H3K9me3 的结合。因此，一个单独的活化标记可拮抗附近抑制因子的结合。

染色质重塑

DNA 上核小体的流动性及组蛋白的驱逐依赖于一类被称为染色质重塑酶的 ATP 依赖性因子（Cairns 2005；Saha et al. 2006a，b）。这类酶在染色质代谢和基因调控中具有不同作用，而且，并不奇怪的是，已经进化出数个家族的该类蛋白质。基于 ATP 酶结构域与 DNA 转位酶的结构相似性，这些家族似乎有相关的机制。然而，它们的功能已变得高度专一化，其中的原因部分是因为结合亚基的差异，部分是因为 ATP 酶结

构域功能的细微差别。目前有 5 个家族：SWI/SNF（BAF）、ISWI（仿制开关）、INO80、NURD/Mi-2-CHD1-9 和 SWR1（表 1.3）。这些酶及其相关亚基利用 ATP 的能量进行滑动或驱逐组蛋白八聚体。八聚体通过大约 120 个特异性的相互作用接触 DNA 的 14 个小沟。同时，破坏这些接触所需求的能量需要 ATP 的水解，对于一个八聚体（K_d约 10^{-10}），曾估计能量需要为 12～14 kcal。

表 1.3 ATP 重塑机器及其亚基

INO80			SWR1		
物种	酵母	人	酵母	人	
复合物	INO80	INO80	SWR1	SRCAP	
同源亚基[a]	Ino80	hIno80	Swr1	SRCAP	
	Rvb1、Rvb2	Tip49a、Tip49b	Rvb1、Rvb2	Tip49a、Tip49b	
	Arp4、5、8，Act1	BAF53a，Arp5、8	Arp4、6，Act1	BAF53、Arp6	
	Taf14		Yaf9	GAS41	
	Ies2	hIes2			
	Ies6	hIes6			
			Swc4/Eaf2	DMAP1	
			Swc2/Vps72	YL-1	
			Bdf1		
			H2AZ，H2B	H2AZ、H2B	
			Swc6/Vps71	ZnF-HIT1	
特有亚基	Les1、Ies3-5、Nhp10	Amida、NFRKB、MCRS1、FLJ90652、FLJ20309	Swc3、5、7		
ISW1 亚家族	ACF/CHRAC				NURF
物种	酵母	酵母	酵母	人	人
复合物	ISW1[a]	ISW1[b]	ISW2[b]	ACF[c]	NURF
同源亚基[a]	Isw1	Isw1	Isw2	hSNF2H	hSNF2L
			Itc1	WCRF180/hACF1	BPTF
特有亚基	Ioc3	Ioc2、Ioc4			RbAP46、RbAP48
Mi-2/CHD 亚家族	CHD1		Mi-2/CHD		
物种	酵母	人	人		
复合物	CHD1	CHD1	NuRD[d]		
同源亚基[a]	Chd1	CHD1	Mi-2α/CHD3、Mi-2β/CHD4		
特有亚基			MBD3；MTA1、2、3；HDAC1、2；RbAp46、48；p66αβ；DOC-1		

续表

SWI/SNF				
物种	酵母	酵母	人	人
复合物	SWI/SNF	RSC	BAF[c]	PBAF
同源亚基[a]	Swi2/Snf2	Sth1	BRG1 或 hBRM	BRG1
	Swi1/Adr6		BAF250/hOSA1	
		Rsc1[e]		BAF180
		Rsc2		
		Rsc4		
	Swi3	Rsc8		BAF155、BAF170
	Swp73	Rsc6	BAF60[a]	BAF60[a]
	Arp7、Arp9	Arp7、Arp9	BAF53	BAF53
	Snf5	Sfh1	hSNF5	hSNF5
			BAF57	BAF57
			β-actin	β-actin
特有亚基	Swp82、Taf14、Snf6、Snf11	Rsc3、5、7、9、10、30，Htl1、Ldb7 和 Rtt102		

经 Elsevier 许可，改编自 Bao 和 Shen（2007b）。

a 第一行表示核心 ATP 酶。

b 除了 Isw2 和 Itc1，变种 ISWI2 复合物也包含 Dls1 和 Dpb4。

c 除了 hSNF2H 和 hACF1，人 CHRAC 也包含 hCHRAC15 和 hCHRAC17。在人的其他 ISWI 复合物中也已发现了 hSNF2H，这些复合物包括 RSF、WICH、B-WICH 和 SNF2h/NuRD/粘连复合物。

d 人 MeCP1 复合物包含 MBD2，而不包含 MBD3；在人 MeCP1 中也发现了 MEP50 和 PRMT5。

e PBAF 复合物包含 BAF180 和 BAF200 而不包含 BAF250。

SWI/SNF 家族包括酿酒酵母的 SWI/SNF 和 RSC，以及被称为梵天相关基因 1 相关因子（BAF）和 PBAF 的两个主要形式的哺乳动物 SWI/SNF（Mohrmann and Verrijzer 2005）。酿酒酵母的 SWI/SNF 和 RSC 包含识别组蛋白尾巴中乙酰化赖氨酸的溴域。这些溴域增加 RSC 和 SWI/SNF 对染色质的结合亲和力，并且提高它们重塑核小体的能力。在所有 SWI/SNF 家族成员中，ATP 酶亚基包含位于靠近蛋白质羧基端的一个溴域。然而，额外的亚基提供额外的溴域。例如，RSC 包含 15 种已知酿酒酵母溴域中的 8 种。大多数结构域是由催化亚基 Sthl 以外的亚基提供的。

哺乳动物有数种 SWI/SNF 酶的变种（Mohrmann and Verrijzer 2005）。ATP 酶被包含在称为梵天（brahma）相关基因 1（Brg1）或 Brm1（哺乳动物梵天）的两种相关蛋白中。哺乳动物包含称为 BAF 和 PBAF 的两种 SWI/SNF 复合体。BAF 利用 Brg1 或 Brm1 作为 ATP 酶，而 PBAF 则利用 Brg1 作为 ATP 酶。BAF 和 PBAF 共用许多辅助亚基，二者不同的是 PBAF 也包含至少一个称为 BAF180 的新亚基，而 BAF 复合物含有称为 BAF250 的特有亚基。BAF180 或聚溴包含 6 个溴域，据认为这些溴域与 RSC

中的各种溴域有些保守，因此认为人 PBAF 与酿酒酵母 RSC 相关。BAF20 包含特有的富含 AT 的互作结构域（ARID），该结构域可能结合 DNA。这样的一个结构域是在 SWI1 中被发现的，SWI1 是酿酒酵母 SWI/SNF 复合物中特有的。因此，认为 BAF 更像 SWI/SNF 而不是 RSC。PBAP 和 BAP 是果蝇（*Drosophila melanogaster*）复合物，类似于人 PBAF 和 BAF。与人不同的是，果蝇只包含单个 ATP 酶，被命名为梵天或 Brm。

注意，所有 ATP 依赖性重塑机器含有肌动蛋白和与肌动蛋白相关的蛋白质，或 ARP。虽然 ARP 似乎不能聚合或水解 ATP，但其具体功能还不清楚（Chen and Shen 2007）。几个研究表明，ARP 亚基在重塑复合物的组装中发挥作用。在许多高等真核生物染色质重塑和修饰复合物中，发现 ARP4 与肌动蛋白呈现化学当量的数量关系。在酿酒酵母中，ARP7 和 ARP9 似乎取代 SWI/SNF 和 RSC 中的 ARP4 和肌动蛋白。曾假设肌动蛋白的聚合可能单独地或黏附于染色质上来控制 SWI/SNF 通过细胞核的运动。例如，研究者可能设想这样一个模式，在该模式中肌动蛋白将附着到 SWI/SNF 上并移动染色质进入转录工厂。

SWI/SNF 家族的另一个有趣的特征是它产生混乱的核小体阵列，并且可将 H2A-H2B 二聚体甚至是八聚体转移到另一个 DNA 分子或体外伴侣上（即驱逐）。全基因组的相关研究已经表明，SWI/SNF 家族的成员可结合重叠基因，然而，遗传分析表明这些酶在体内具有不重叠的作用。SWI/SNF 家族的成员涉及体内基因活化，并在生化系统中可促进 Pol II 的延伸。SWI/SNF 家族的成员也可直接结合某些激活因子（Peterson and Workman 2000）。

SWI/SNF 和 RSC 的 EM 结构已经被解析，从而揭示出一个在其中可塑造核小体的中央腔模型。该结果表明，作为重塑机制的一部分，ATP 酶复合物可以包封核小体（Smith et al. 2003；Leschziner et al. 2007）。目前的模型也显示 RSC 似乎能完全包封核小体（图 1.5），而 SWI/SNF 只与沿核小体一个面上的一圈 DNA 相互作用（Dechassa et al. 2008）。人 SWI/SNF 模型酷似 RSC 模型。多种数据表明这些酶直接从核小体二分体接触 DNA 的两个转角，并且将 DNA 拖向二分体，再将 DNA 拖向组蛋白八聚体周围，最后将 DNA 从另一个末端拖出来，从而有效地移动八聚体在 DNA 上的位置（Saha et al. 2006a）。已提出了一个称为“波-棘轮-波”模型的详细模型（Saha et al. 2006a）。虽然才刚刚开始调查溴域-乙酰基赖氨酸相互作用对重塑机制的影响，但早期的数据表明修饰不仅增加亲和力，也改变重塑产物的分布，并可改变酶将八聚体转移到另一 DNA 分子上的能力（Ferreira et al. 2007）。

SWI2 和 STH1 分别是 SWI/SNF 和 RSC 的催化亚基，与这两个亚基不同的是，ISWI 染色质重塑酶享有不含溴域的 ISWI ATP 酶。然而，ISWI ATP 酶的确含有 SANT 和 SLIDE 结构域，这些结构域有助于将酶靶向于染色质（Mellor and Morillon 2004；Saha et al. 2006b）。在酿酒酵母和哺乳动物中，有两种 ISWI 同系物（酿酒酵母中的 Isw1 和 Isw2）。ISWI 在不同复合物中与不同亚基结合，包括 RSF、NURF、CHRAC 和 ACF。在 NURF 和 ACF 中，分别在 ACF 和 NURF301 亚基中发现了溴域。不像 SWI/SNF，ISWI 家族似乎在体外产生规则间隔的或有序排列的核小体阵列。在酵

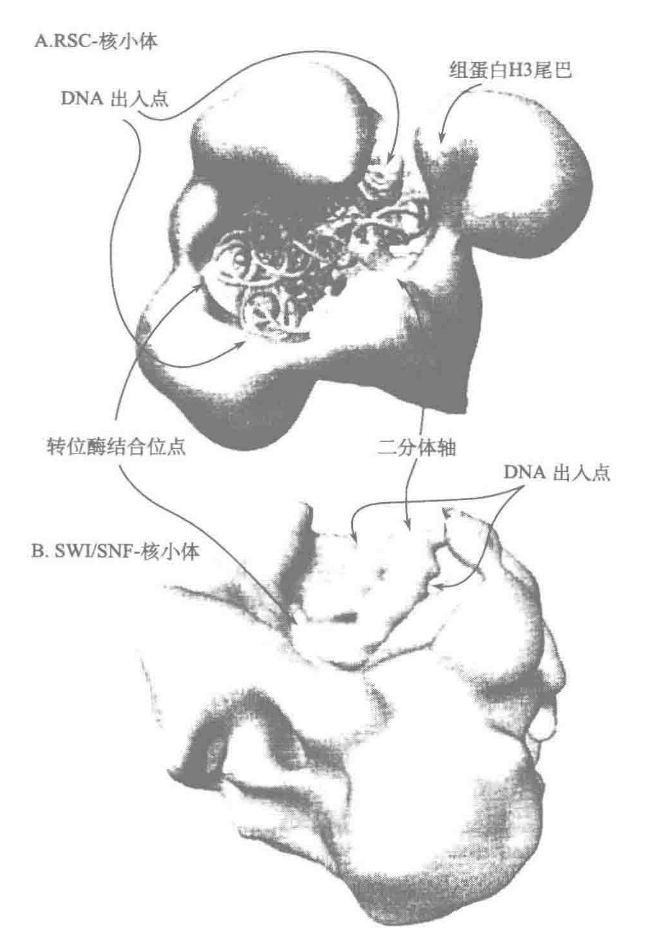


图 1.5　RSC-核小体（A）和 SWI/SNF-核小体（B）EM 结构的模型。在 RSC 中，ATP 酶包封核小体，而在 SWI/SNF 中，ATP 酶部分包封核小体。图中显示了 DNA 出入点的位置、核小体二分体和 DNA 转位酶的接触点（RSC-核小体模型图改编自 Leschziner et al. 2007，© National Academy of Sciences，U. S. A.；经 American Society for Microbiology 许可，SWI/SNF-核小体模型图改编自 Dechassa 2008）。

母中，Isw2 ATP 酶复合物涉及抑制，并且是抑制因子的靶标，如抑制因子 Ume6 及其伴随的 Rpd/Sin3 复合物（下文进一步描述）(Fazzio et al. 2001)。此外，Isw2 ATP 酶复合物对 Ssn6/Tup1 共抑制因子的核小体的定位功能也是必需的（Fazzio et al. 2001）。有趣的是，抑制从某些酵母基因的 3′ 端发出的反义转录也需要 Isw2 复合物（Whitehouse et al. 2007）。在酿酒酵母中，ISWI 也涉及转录的延伸（Mellor and Morillon 2004）。

在哺乳动物中，重塑蛋白的 CHD 或 Mi-2 家族由 9 个家族成员组成（Hall and

Georgel 2007)。借助直接位于 SWI/SNF 样解旋酶结构域氨基末端的双重染色域，可表征这些成员，该染色域似乎识别具有高亲和力的 H3K4me3。据认为，CDH1 是 SAGA 的组分，且在酿酒酵母中与 SLIK HAT 有关，而且在基因活化中与这些 HAT 协作。在体外，CHD1 滑动核小体，并支持有序阵列的组装。在体外，CHD1 也可以从自由的八聚体和伴侣组装核小体，并且已报道在体内基因活化期间 CHD1 插入 H3.3 (Konev et al. 2007)。CHD3 (Mi-2a) 和 CHD4 (Mi-2b) 是 NuRD 复合物的 ATP 酶亚基，并且与 RbAP46/48 一起结合 HDAC1 和 HDAC2。NuRD 涉及基因沉默，并且在体外同数种序列特异的抑制因子相互作用。CHD7 与果蝇的天命 (Kismet) 蛋白相同，天命蛋白广泛联系着基因活化和转录延伸。天命蛋白属于三胸因子，三胸因子对果蝇发育期间转录的激活状态的维持非常重要。三胸蛋白组也包括 SWI/SNF 和 MLL/SET1 复合物。已知 CHD8 与 CTCF 结合，可能参与边界功能（描述如下）。除了天命蛋白，CHD 家族成员一般与许多辅助亚基相结合。

从酿酒酵母到人的有机体中均发现了 INO80 复合物 (Bao and Shen 2007a)。INO80 非同寻常，因为它的 ATP 酶结构域显示一个大的插入物，该插入物在大小上比其他 ATP 酶中发现的插入物更大。由于在所谓的分裂 ATP 酶结构域中的相似性，INO80 最初被归入 SWR1，但相对于 INO80，SWR1 具有非同寻常的一些特性，因此最近 INO80 又被重新分类，不再归入 SWR1。在哺乳动物细胞中，INO80 与序列特异性蛋白 YY1 紧密结合，这有助于在体内将 INO80 靶向于其结合位点 (Cai et al. 2007)，其在体外也滑动核小体。不像 INO80，SWR1 能够特异地将 H2AZ 转移进带有 H2A 的核小体 (Korber and Horz 2004)。遗传学研究也表明 *Swr1* 和 *Htz* (*H2AZ* 基因) 突变体有相似的表型。另外，基因表达和 ChIP 研究显示，SWR1 在 H2AZ 插入的地方调控基因 (Bao and Shen 2007a)。

通用转录机器

哺乳动物基因调控涉及激活因子、抑制因子、通用转录机器和染色质间的相互作用。通用转录机器由 RNA Pol II，通用转录因子 (GTF) TFIIA、TFIIB、TFIID、TFIIE、TFIIF、TFIIH，以及被称为调解因子的一种主要共激活因子组成，调解因子将激活因子连接到 RNA Pol II 和 GTF 上（见第 12 章中的表 12.2～表 12.4，该章描述了 TFIID、调解因子和 GTF）(Hahn 2004; Roeder 2005; Kornberg 2007)。TFIID 是一个多亚基蛋白，包含 TATA 盒结合蛋白 (TBP) 和 TBP 相关因子或 TAF (Albright and Tjian 2000)。正如下文所讨论的，虽然 TAF 的主要作用可能是识别不同的核心启动子元件，但是 TAF 具有共激活因子功能。从晶体学或 EM 研究中已经获悉了 GTF、调解因子、Pol II 和 TFIID 的结构。RNA Pol II 的一个重要特征是七肽重复，该重复构成了最大的 RNA Pol II 亚基的羧基末端结构域 (CTD)，即 Rpb1。该七肽 ($Y_1S_2P_3T_4S_5P_6S_7$) 在酿酒酵母中重复 26 次，而在哺乳动物中重复 52 次。该 CTD 被参与转录调控的不同激酶广泛磷酸化，并且这些事件调控起始、延伸及各种 mRNA 加工事件（下文讨论）。生化研究表明，GTF 支持基础转录，并且执行很多转录起始所需的催化功能。诸如调解因子的共激活因子和 TFIID 的 TAF 亚基是激活因子和 GTF 间

的桥梁。应该指出，调解因子在体外也增强基础转录。

在真核生物中也发现了 TBP 相关因子，即 TRF 的一个家族，并且在某些环境中 TRF 可代替 TFIID (Reina and Hernandez 2007; Torres-Padilla and Tora 2007)。在果蝇中发现了 TRF1，在所有的后生动物中发现了 TRF2，而在脊椎动物中发现了 TRF3。TRF3 与 TAF3、TFIIA 和 TFIIB 相联系，而且能结合含有 TATA 的启动子并支持体内外的转录。据认为，TRF3/TAF3 复合物取代 TFIID，并且支持参与肌形成和其他可能过程的基因的活化。除了肌形成基因，已经获知各种 TRF 因子的数个体内靶点 (Reina and Hernandez 2007)。

大多数 GTF、TAF 和调解因子亚基有多种称为异构体的拼接变体，当考虑到异构体的所有组合的可能性时，通用机器的实际蛋白质成分会有相当大的变化。因此，鉴于读者已经了解包含不同异构体的 PIC 可能平行地存在于相同的细胞中，本章中我们通篇使用“通用机器”这一术语。

调控区的组织

典型的哺乳动物基因可被细分为有助于转录活性的数个区域，包括核心启动子、近端启动子和增强子（图 1.6）(Smale and Kadonaga 2003; Maston et al. 2006; Yang et al. 2007)，其他复合元件（如 LCR）以时间或细胞特异的方式调节各组相关基因 (Dean 2006)。据认为，增强子和近端启动子连同其他 DNA 元件形成称为“活性染色质中心”或 ACH 的高级结构 (de Laat and Grosveld 2003)。也可在含有 Pol II 和通用转录因子的致密核点 (nuclear foci) 中发现进行类似调控的基因，这些核点被称为“转

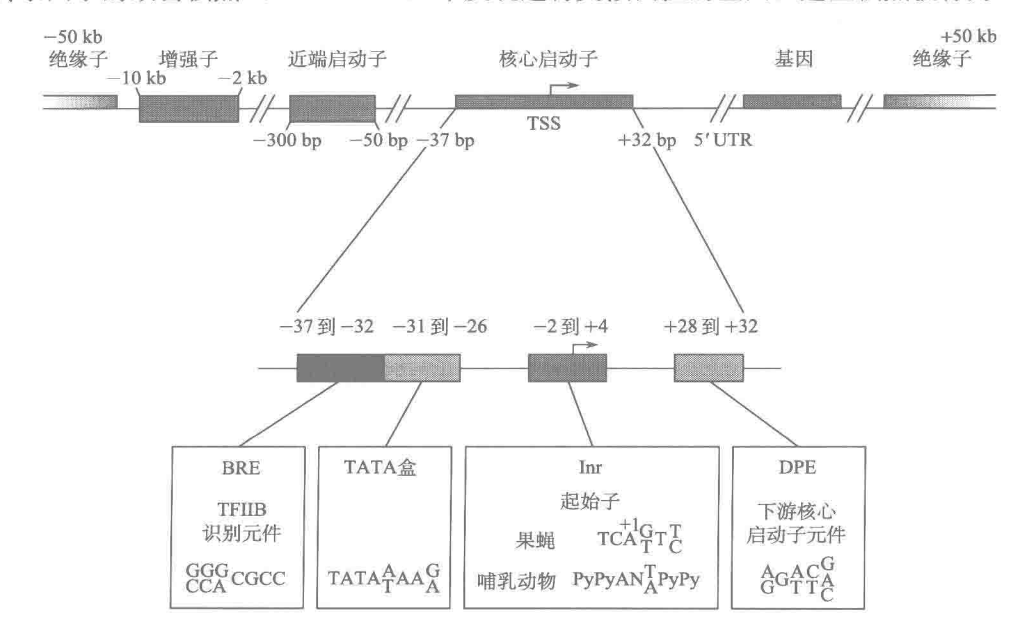


图 1.6 一个调控基因的启动子区。图中显示了各种核心启动子和上游启动子元件。注意，增强子、LCR 和绝缘子的位置在基因间有相当大的变化（底部：经许可改编自 Smale and Kadomaga 2003，© Annual Reviews，www.annualreviews.org）。

录工厂”（Sexton et al. 2007；Carter et al 2008）。在成纤维细胞中有多达 2000 个这样的工厂，而在较小的原代细胞中大约有 200 个转录工厂，转录工厂的数量似乎与细胞核体积有关。活性基因或共线性的基因集合两侧有边界元件，据认为这些元件拴在核基质或者诸如核仁的其他核结构上（Zhao and Dean 2005；Valenzuela and Kamakaka 2006；Dorman et al. 2007）。因此，当考虑基因是怎样被调控时，应考虑很多因素，而不仅仅考虑一个基因或基因座的简单线性组织。我们简要阐述一下构成一个被调控基因座位的各种元件，并讨论一下与每种元件相联系的染色质修饰的类型。

核心启动子

核心启动子结合 Pol II 和通用转录机器，并且位于 TSS 任一侧 35～40 bp 处（Smale and Kadonaga 2003；Maston et al 2006），可细分成许多不同的序列元件。这些元件包括识别通用因子 TFIIB 的 IIB 识别元件（BRE）、识别 TFIID 中 TBP 亚基（下文详细描述）的 TATA 盒、识别 TFIID 中 TAF1 和 TAF2 并包围转录起始位点的起始元件（Inr），以及在果蝇中更为普遍的识别 TFIID 中 TAF6 和 TAF9 的下游启动子元件（DPE）（下文讨论）。在某些启动子中还发现了其他下游元件，包括下游核心元件及模体十元件（MTE）（Maston et al. 2006）。

这些元件中只有一部分出现在任何个体的启动子中。例如，基于哺乳动物中注释的 TSS 的位置，只有约 25%的哺乳动物基因含有一致的 TATA 或者与其类似的富含 AT 的序列（Yang et al. 2007）。目前的看法是 TATA 盒一般位于被调节的基因中，如细胞或发育特异的基因，或对信号作出应答的基因。绝大多数所谓的广泛表达的“持家”基因（>60%）都缺乏 TATA 盒；大多数这些基因在 TSS 一定范围的上游含有 0.5～2 kb的非甲基化 CG 富集区（CpG 岛），该 CG 富集区位于约 100 bp 窗口内。尽管持家基因倾向于含有 CpG 富集区，但某些受调控基因也富含 CpG。与 TATA 盒相比，很大比例的基因含有一个 Inr 元件，但该数量仍小于 50%，更少的基因含有 DPE 或 BRE（见 Yang et al. 2007 及其中的参考文献）。这些基因是否以通过晶体结构（见下文）所预测的方式需要或结合 TFIID，仍有待确定。全基因组研究已开始揭示启动子和增强子组织的某些通用原则。

全基因组 DNase I 高敏感性分析表明，22%的高敏感性位点在细胞类型间是保守的（在涵盖 6 种不同细胞系的有限研究中）（Xi et al. 2007）。然而，正如通过表达谱数据及注释的 TSS 的存在所确定的，在这些保守的高敏感性位点中，86%是富含 CpG 的启动子。这些研究也表明，在缺失 TATA 盒的情况下，通过牵扯 TFIID，TAF1 被结合到显著的一部分富含 CpG 的启动子上（Kim et al. 2005）。在同一研究组进行的一项单独研究中，基于 TAF1 和 Pol II 的结合，以及 H3、H3K9/K14 去乙酰化、H4K5/8/12/16 乙酰化、H3K4me、H3K4me2 和 H3K4me3 的存在，在 ENCODE 阵列上的全基因组 ChIP 研究鉴定了启动子（见下文）（Heintzman et al. 2007）。在各种修饰模式的基础上，这些作者鉴定出具有增加表达水平的四类启动子（P1～P4），平均起来，这四类启动子与 TAF1、Pol II，以及 H3/H4 乙酰化的数量相关（Kim et al. 2005）。例如，在酿酒酵母中，H3K4me3 的水平在启动子处达到顶峰，而 H3K4me2 和 H3K4mel 的水平在 TSS 下游有所提高。在 TSS 周围，乙酰化呈现双峰分布，并且 H3 在 TSS 处普

遍缺失，从而表明 TSS 处的核小体被去除，并且两侧的核小体被乙酰化。缺失的数量与基于转录物丰度的启动子的活性相关。这些数据同单个启动子的早期 ChIP 研究结果一致。因此，这些作者推断超乙酰化、H3K4me3、TAF1 及 Pol II 是活性启动子的标志。然而，要注意在发表的研究中有点不清楚的是，H3 缺失与否主要是在 CpG 岛启动子上被发现的，还是在所有启动子上被发现的。

一项不同的研究鉴定了所谓的“共同修饰模块”，该模块由 17 个不同的染色质标记构成，而这些染色质标记发现存在于 CD4 T 细胞中 25%的活跃启动子上（Wang et al. 2008）。一部分标记发现分布在 TSS 的上游，而与此不同的另一部分标记分布在 TSS 的下游。这些作者将他们的数据细分成从低到高的三种表达类型（I、II 和 III）。类型 I 富集于 H3K27me3 和 H2AZ 等，但没有乙酰化，而且类型 I 包括参与发育和分化的基因。类型 II 由 H3K36me3 和共用修饰模块组成。类型 III 包括共同修饰模块，外加 H4K16ac、H4K20mel 和 H3K79mel/2/3，它似乎主要界定持家基因。高分辨率作图显示，很多这些修饰似乎共定位到相同的位置，但还不清楚这代表一种群体效应，还是这些修饰确实被共同定位到单一核小体上。这些修饰中某些修饰的作用在上下文中都进行了讨论，尽管大多数修饰的特异性作用仍有待研究。已经提出，参与相同通路的共调控基因有相似的组蛋白修饰模式（Kurdistani et al. 2004）。

近端启动子区

紧挨着核心启动子的上游，通常在第一个 250～300 bp 内，是一簇序列特异性转录因子的结合位点，该区域被称为近端启动子（Maston et al. 2006）。一般通过检查核心启动子和 TSS 上游系统发育上保守的序列，可确定近端启动子的边界（见第 5 章和第 6 章）。近端启动子可与结合在远端增强子上的某些相同因子结合，或者与特有的因子结合。对个别基因的很多研究已表明，近端启动子自身具有低的活性水平，但这种活性可被远端增强子极大地提高。相反，在缺失近端启动子的情况下，远端增强子没有功能。很有可能增强子和近端启动子共用某些共同的共激活因子，如调解因子；调解因子使二者之间形成环并将增强子带到通用机器附近。由于定位远端增强子及将它们整合进叠连芯片阵列的不确定性，全基因组 ChIP-chip 研究一般检查近端启动子及 TSS 上游或下游固定数量的序列。因此，在正在进行的序列特异性转录因子的很多研究中，要注意这些研究一般检查近端启动子，而 ChIP-Seq 没有该局限性。个别基因的研究表明，在通用机器向核心启动子的募集中，近端启动子发挥关键作用。

增强子

增强子可位于基因上游数千到数万个碱基处、内含子内，偶尔出现在编码区 3′ 端之后（de Laat and Grosveld 2003；Zhao and Dean 2005；Maston et al. 2006）。增强子结合序列特异性激活因子的组合体，有时这种结合以协作方式发生，这些结合的结果是形成增强体（见第 11 章）。很多基因含有结合不同因子的多个增强子，这些增强子使基因在不同类型的细胞中或在对单一信号的应答中被组合性地激活。正如上文所描述的，增强子结合近端启动子形成活性染色质中心。有数个模型描述了增强子如何与近端启动子联系。增强子需要近端启动子才能起作用，这一结果可能是这种相互作用的中央面（central facet），并且很可能是活性染色质中心的组织原理。有很多不与附近的非靶基

因启动子联系的增强子的例子。原因之一是很多激活因子经历特异性的、协作性的蛋白质-蛋白质接触，这些接触可能促进增强子-启动子的相互作用。因此，近端启动子可能结合与增强子不相容的序列特异性因子。另外，桥接增强子和近端启动子的共激活因子可能只与激活因子的特定构型相容。而且，通过成环/拴系（见上文）及受限制的互作结构域的形成，绝缘子元件调控增强子和近端启动子的相互作用。

关于远端增强子如何与近端启动子联系以协作性地募集通用机器，只有一些有限的了解。例如，3C研究已表明在基因活化期间，雄激素受体（AR）从距前列腺特异性抗原（PSA）基因 *TSS* 4.2 kb的增强子成环到近端启动子上（Wang et al. 2005）。很显然，近端启动子也含有AR结合位点。另外，有关PSA的ChIP研究表明，Pol II和CBP均定位于启动子和增强子位点上。对于该观察结果一个较为合理的解释是Pol II与增强子交联，因为增强子在Pol II结合的启动子的附近。在其他情况下，似乎存在一个便利的跟踪机制，该机制中染色质被单向乙酰化，并且增强子复合物似乎沿着乙酰化染色质移动，直到到达启动子（即HNF4α，人β-珠蛋白）（Hatzis and Talianidis 2002；Zhao and Dean 2005）。在其他例子中，这种远程跟踪需要Pol II和TBP及基因间的转录，以促进增强子复合物追踪到启动子上。这类追踪的一个很好的例子是β-珠蛋白基因座（Zhao and Dean 2005）。全基因组研究通常未发现基因的增强子和启动子间乙酰化的长踪迹，相反，却显示出分离的超乙酰化簇，从而表明跟踪可能限于少数增强子-启动子相互作用（Roh et al. 2007）。当3C技术同机制方法相结合时，增强子-启动子联系的问题就可能变得更清楚。

全基因组研究似乎不支持与增强子相关的组蛋白修饰的类型。所有已经被单独研究和全基因组研究的增强子一般都显示DNase I高敏感性和超乙酰化，然而，全基因研究表明增强子更复杂。一项研究表明，增强子与DNase高敏感位点、p300的富集及其相关修饰，以及H3的普遍缺失相关。和启动子相比，增强子显示富集于H3K4me1而不是H3K4me3中（Heintzman et al. 2007）。在某些增强子上发现了TAF1、Med1和RNA Pol II，但是富集水平不及在启动子上发现的富集水平。另一研究认为增强子对应于DNase I高敏感性，并被严重乙酰化，但是作者主张有多种与增强子相关的不同的修饰模式，其中某些修饰模式富集在H3K4me2和H3K4me3中（Barski et al. 2007）。尤其是就H3K4甲基化水平而言，这种不一致性表明不同的增强子可能使用不同的作用机制，这种不一致性可能还反映出在体内全基因组水平上鉴定增强子的难度。

基因座控制区

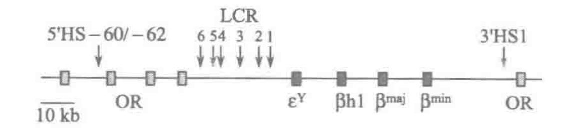
LCR是多种增强子、绝缘子和其他元件的组合，它以时间或细胞特异性的方式控制各组基因（Dean 2006；Mahajan et al. 2007）。LCR包含多个不同DNase I高敏感位点。借助于对染色质结构的影响，LCR通过一段延长的DNA区域调节转录，而且，正如DNase I高敏感性所测定的，LCR似乎诱导并维持该区域对转录因子的可及性。LCR是在研究转基因鼠β-珠蛋白基因座的过程中鉴定的。β-珠蛋白转基因在低水平表达，并且该表达水平受到插入染色体的位点的强烈影响。然而，当该转基因含有来自珠蛋白基因座远端的特定DNA片段时，则观察到高水平表达，而且这种高水平表达与位置无关（或与整合位点无关），现已经知道该DNA片段包含LCR。

LCR 的界定特征是：①在转基因实验中，给其连接的基因赋予高水平诱导表达的能力；②表达水平直接与拷贝数关联；③在染色质环境中增强转录的能力；④在长距离上打开染色质的能力；⑤赋予位置无关性表达的能力，即使在着丝粒异染色质存在的情况下也是如此，而着丝粒异染色质一般导致被整合的转基因失活。

一般而言，LCR 与细胞特异性增强子有共同的特性，因为它们都与 DNase I 高敏感位点一致，并都与典型的转录因子结合，而且，它们可以刺激周围染色质的超乙酰化。尽管有这些相似之处，但 LCR 与增强子有明显的区别，因为在许多情况下，只有在整合进染色体中时，LCR 才增强转录。此外，在转基因小鼠中，典型的增强子不能单独给予高水平的位置无关性表达。

基于 3C 和 RNA 陷阱（RNA Trap）研究，目前的看法是 LCR 环状突出并与其连锁基因的近端启动子及增强子相互作用。据报道，LCR 和连锁基因间接触的频率及持续时间控制基因具有转录活性的时段。LCR 的各种缺失仍然保留最小的功能，但它们会降低基因表达的活性和位置独立性。

LCR 的范例是 β-珠蛋白基因座，已在多种生物中对该基因座 LCR 进行了广泛研究（图 1.7）。在鼠中，β-珠蛋白基因座由 4 个基因组成，分别被称为 ε、βh1、大 β（β major）和小 β（β minor）。LCR 分别位于 ε 和大 β 基因远端 6～60 kb 处。ε 和 βh1 基因在胚胎中表达，而大 β 和小 β 在成体中表达。8 个高敏感（HS）位点中，3 个是组成型的，并出现在前红系细胞中（5′ HS-60/-62、HS5 和 3′ HS1），5 个是红系特异性位点（HS1、2、3、4 和 6）。正如 3C 技术所鉴定的，在前红系细胞中，组成型 HS 间似乎相



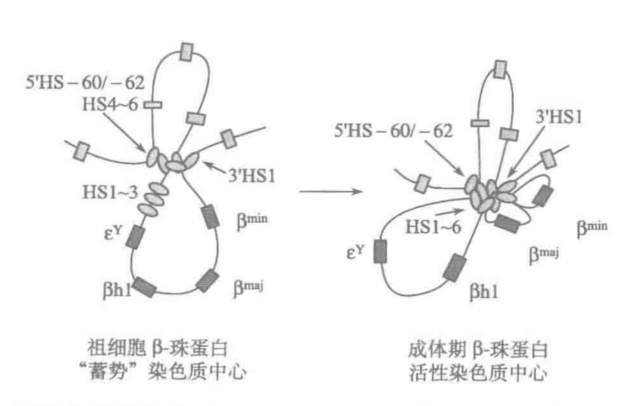


图 1.7　LCR 和活性染色质中心（经 Elsevier 许可，改编自 Dean 2006）。图中显示了鼠 LCR 和珠蛋白基因座的组织。图中还显示了各种元件折叠进前红系细胞中的蓄势染色质中心和成熟红系细胞中的活性染色质中心。各种环都有 3C 证据。图中也显示了 DNase I 高敏感性位点的位置。

互作用，从而形成蓄势染色质中心（poised chromatin hub)。在胚胎细胞中，HS1-3 和胚胎珠蛋白基因加入以形成活性染色质中心；在成体细胞中，胚胎基因环化突出，并且成体基因加入活性染色质中心。包括 GATA-1、EKLF 和 FOG-1 的红系特异性转录因子和共激活因子，对于活性染色质中心的稳定是必需的（见第 10 章)。在鼠中，附着在活性中心的活性基因被严重地乙酰化，但是无活性基因却没有；在人中，好像有胚胎细胞的持续的乙酰化。

在后生动物基因组中已经表征了大量的 LCR，并且 LCR 组织的多样性表明 LCR 以多种方式调控靶基因（Dean 2006)。有很多表征得非常清楚的 LCR，其中不乏一些激动人心的例子，如生长激素 LCR，该 LCR 差异性地控制一簇基因。这些基因的其中之一在腺垂体中表达，而其他基因在胎盘中表达。Th2 LCR 在 T 细胞分化期间控制白细胞介素 *Il-4*、*Il -5* 和 *Il -13* 基因的表达。该 LCR 实际上定位在 *Il-5* 和 *Il-13* 之间的 *Rad50* 基因内。最近的数据也表明细胞因子基因在由 SATB1 蛋白界定的高级染色体结构域中发挥作用，SATB1 蛋白通过结合碱基非配对区，控制细胞因子基因的表达（Galande et al. 2007)。

绝缘子

绝缘子 DNA 元件及其伴随的结合蛋白在转录的调控中发挥多重作用（Zhao and Dean 2005；Valenzuela and Kamakaka 2006；Dorman et al. 2007；Maeda and Karch 2007；Wallace and Felsenfeld 2007)。绝缘子的两个主要作用是调控异染色质向转录活性区域的扩展，以及通过阻止增强子作用于其非预期靶标的启动子，以维持增强子的基因特异性。例如，哺乳动物中 β-珠蛋白基因位于两簇气味受体基因之间。因为这两类基因不在相同的组织中共表达，必须有一种准备就绪的机制，来阻止其中一类基因的增强子和启动子与另一类基因的增强子和启动子发生相互作用。显然，绝缘子阻止这种串扰。然而，另一个要考虑的是，在红系细胞中缺乏结合气味受体基因近端启动子的因子，在感觉神经元中缺乏红系因子。最后，由于人类基因组的继续分析，已经开始出现许多例子，在这些例子中，一个基因的调控元件实际上出现在另一个基因中，或被另一个基因隔开（如上文描述的细胞因子 LCR)。

正如上文所讨论的，绝缘子以其最简单的形式发挥两个作用，通过与它们结合的蛋白质，时常能辨别它们的功能，但这一点并不总是能做到。在第一个作用中，绝缘子阻碍异染色质接近基因组的活性区，该功能通常被称为屏障功能。在第二个作用中，绝缘子阻碍增强子与错误启动子进行不恰当的联系，该功能通常被称为增强子封闭。绝缘子的相关文献非常复杂，并且绝缘子的功能与高级染色质组织以及细胞架构有关。

屏障绝缘子通常与序列特异的转录因子、染色质修饰及重塑相关。增强子封闭绝缘子显示出倾向于形成环且有时附着于如核仁和核纤层这样的核结构上。脊椎动物 β-珠蛋白绝缘子的 5′ HS4 元件含有上游刺激因子（USF）和 CCCTC-结合因子（CTCF）的结合位点。USF 是个募集染色质乙酰化及重塑蛋白的转录激活因子。CTCF 是个 11 锌指 DNA 结合蛋白，该蛋白质可识别一个 20 bp 的富含 GC 的位点，该位点具有来自全基因组分析的一个已知一致序列。CTCF 似乎是介导增强子封闭的主要脊椎动物蛋白。在孤立的情况下，USF 显示出清楚的屏障作用，而 CTCF 只显示增强子封闭活性。绝缘

子的另一作用可能是通过形成染色质环的立体定向排布而组织活性染色质中心，染色质环的立体定向排布将活性及无活性元件定位到不同子域。正如通过 3C 所确定的，β-珠蛋白活性染色质中心中的接触点与 CTCF 的位置最相关，而不是与 DNase I HS 位点的边缘最相关。

全基因组研究已表明，细胞中 10%的 DNase I 高敏感性位点定位于 CTCF 结合位点上（Barski et al. 2007；Kim et al. 2007；Xi et al. 2007）。最近的全基因研究也表明，CTCF 与粘连蛋白共定位并且对于粘连蛋白的结合是必需的（Pareiho et al. 2008）。如上文所提到的，CTCF 结合位点时常但不总是在单个基因的两侧，并且在很多情况下，它们位于诸如气味受体等相关基因簇的两侧。通常，CTCF 位点定位在基因组的基因富集区，它们倾向于被大约 20 个富含 H2AZ 的定位核小体包围（Fu et al. 2008）。CTCF 结合位点到 TSS 的平均距离大约为 50 kb。CTCF 似乎与定位在核仁的核磷蛋白和核基质相互作用。因为这些联系，已提出拴系到一种结构上对于 CTCF 的活性很重要。然而，免疫荧光研究已表明 CTCF 在细胞核中分布位置广泛，这些位置并不局限于核仁或细胞核周边。在果蝇（*D. melonogaster*）中，很多被毛翅[Su（Hw）] 形状抑制蛋白（suppressor）结合的绝缘子聚集在核周边，称为绝缘体（Maeda and Karch 2007）。推测绝缘体控制特定转录活跃和不活跃区域的形成。

已经在印迹的 *H19*/*Igf2* 基因座上检查了 CTCF 的绝缘子功能（图 1.8）（Zhao and Dean 2005；Valenzuela and Kamakaka 2006；Wallace and Felsenfeld 2007）。在这种情况下，定位在 *H19* 下游的增强子可控制 *H19* 或位于 *H19* 上游的 *Igf2* 基因。被称为差异甲基化区（DMR）的边界元件位于 *Igf2* 和 *H19* 之间。当 DMR 未甲基化时，就如在母系等位基因中那样，CTCF 结合于 *Igf2* 上的增强子并阻止其发挥作用，这就

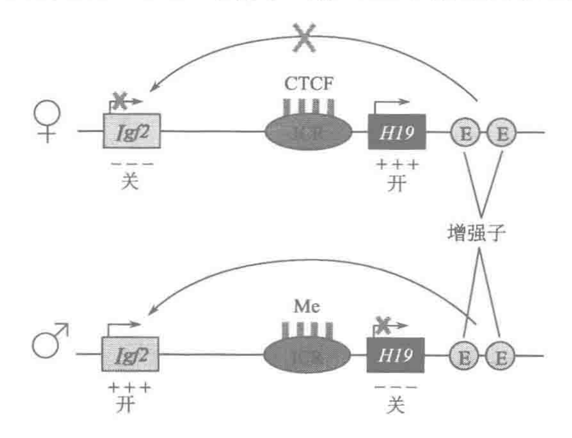


图 1.8 *H19*/*Igf2* 基因座上的 CTCF 绝缘子。*H19*/*Igf2* 是一个印迹基因座的例子，在该基因座中，一个等位基因是母系遗传的并有活性，而另一个等位基因是父系遗传的。在母系等位基因中，CTCF 结合到 DMR 上，并将增强子限定于 *H19*。在父系等位基因中，DMR 被甲基化，这会抑制 CTCF 的结合。此时该增强子环化并激活 *Igf2*（经 Elsevier 许可，改编自 Wallace and Felsenfeld 2007）。

限定该增强子仅对 *H19* 起作用。在父系等位基因中，DMR 被甲基化，从而阻止了 CTCF 的结合并使该增强子激活 *Igf2*。3C 研究已经表明，CTCF 的出现使不同环得以形成；在它存在的情况下，*Igf2* 环化突出，而在它缺失的情况下，*Igf2* 与增强子形成一个不同的环。

在裂殖酵母中，Pol III 转录因子 TFIIIC 也可充当屏障元件，它形成大约 10 个核质体，在那里各种染色质环附着于核周边（Valenzuela and Kamakaka 2006）。已经将这些核质体提议为 TFIIIC 的屏障活性的组织位点，并可能与果蝇中的 Su（Hw）绝缘体相关。

关于果蝇中的绝缘子及其结合蛋白有广泛的文献，也有很多测试屏障绝缘子和封闭绝缘子如何工作的精美的机制性实验。有关该问题的最近的综述和更全面的论述，见 Dorman 等（2007）及 Maeda 和 Karch（2007）的相关文献。

基础转录复合物的组装和起始

纯化的 GTF 和 Pol II 在体外核心启动子上介导低水平的基础转录，但是在缺失共激活因子的情况下，它们不支持激活转录（见下文）。基本转录机器的研究开始于共激活因子发现之前，而且重要的是，在体内可能没有类似基础转录这样的现象。尽管如此，体外基础反应的表征已揭示出发生在 Pol II 起始和从启动子逃离期间的催化事件的很多信息。在大多数关于前起始复合物组装的最初研究中，TBP 用于代替 TFIID，因为 TBP 很小，在其他 GTF 存在的情况下 TBF 可支持基础转录，并且 TBF 可经受凝胶迁移实验（EMSA），以用于详细的结构-功能分析。此外，TFIID 在那时还没有被充分纯化，不足以分析它在前起始复合物（PIC）形成中的作用。

这些早期研究表明，纯化的 GTF 和 TBP 在 DNA 上以逐步的方式组装成转录 PIC（Hahn 2004；Kornberg 2007）。通过 TBP 在 TATA 盒上的结合，使复合物成核，TBP 与 TATA 盒的结合反应需要 TFIIA 或 TFIIB 的辅助，在 TBP 之后，TFIIA 或 TFIIB 可按任意次序结合。现已解析出了 TBP、TBP/TFIIB 和 TBP/TFIIA 与 DNA 复合物的晶体结构，从而更好地揭示出对启动子的识别过程（有关结合的详细结构及信息，见第 10 章）。TFIIA 和 TFIIB 均与 DNA 和 TBP 接触，从而增加了 TBP 结合 TATA 盒的稳定性。在 TFIIB 结合 TBP 后，与 Pol II 关联的 TFIIF 复合物被募集，随后 TFIIE 和 TFIIH 相继结合。

随着单独的 Pol II、延伸的 Pol II、Pol II 与 TFIIS 复合物，以及 Pol II 与 TFIIB 氨基末端结构域复合物的一系列晶体结构的解析，有关起始和延伸的机制出现了令人振奋的认识（Kornberg 2007；Cramer et al. 2008）。在酿酒酵母中，Pol II 由总分子质量为 515 kDa 的 12 个亚基组成，最大的两个亚基——Rpbl 和 Rpb2，形成催化裂缝，并在 Pol I、Pol III 及原核 RNA 聚合酶的 β′ 和 β 亚基中有同源蛋白。这些亚基在转录起始和催化中发挥一般作用（Hahn 2004）。主要由 Rpb1 和 Rbp2 形成的大的中央裂缝（central cleft）（Hahn 2004；Kornberg 2007）具有三个主要特征，分别称之为夹钳（clamp）、下颌/叶（jaw/lobe）和内壁（wall）。夹钳极易移动，并在缺乏 DNA 时呈开放构型，但在有 DNA 时则向下夹住。DNA 通过裂缝进入并行进到内壁底部附近的催

化位点中（图 1.9）。催化位点含有一个镁离子，它对磷酸二酯键的形成很重要。两个通道可使核苷酸进入催化位点，并且可使 RNA 离开催化位点。称为“桥梁螺旋”的 α 螺旋横跨裂缝，并转运 DNA 通过催化位点。关于催化的机制及催化结构域中各种移动元件的运动已有较多的研究成果。有关延伸过程的分子细节，见 Kornberg（2007）和 Cramer 等（2008）。

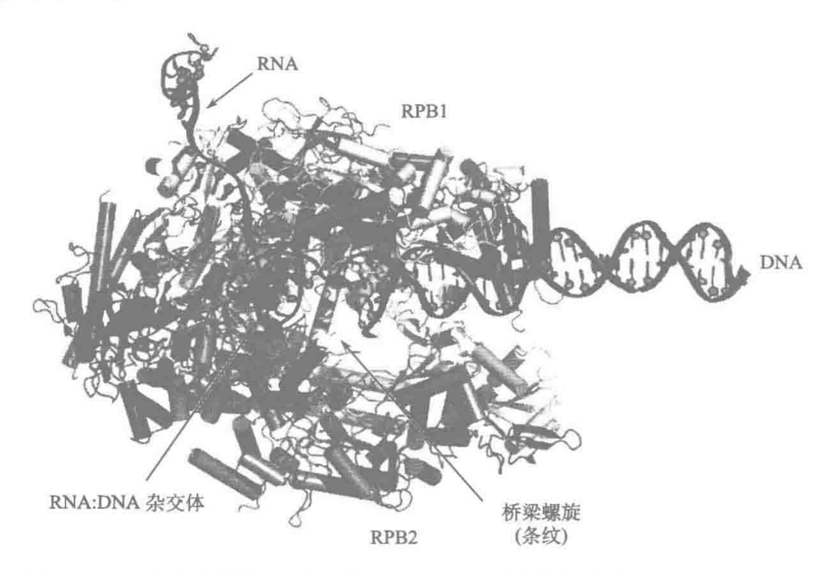


图 1.9　RNA Pol II 延伸复合物。用 PyMOL 绘制的来自（PDB file 1R9S）的酿酒酵母 Pol II 延伸复合物。对来自极端嗜热菌结构（PDF，2O5I）的 DNA-RNA 杂交体建模，取代来自酵母结构的最初 RNA -DNA 杂交体。正在进入的 DNA 和正在出去的 RNA 被人为延长，以更好地说明这些路径。颜色标记：Rpb1（浅灰；顶部）；Rbp2（中等灰；底部）；Pol II 的残余部分（深灰；边缘）。对 DNA 和正在出来的 RNA 进行了标记，并呈非常深的灰色。桥梁螺旋以条纹作标志（由 J. Heiss 绘制，UCLA）。

TSS 进入到活性裂缝的定位是 TBP、TFIIB 和 Pol II 间一系列蛋白质-蛋白质和蛋白质-DNA 相互作用的结果（Chen and Hahn 2003，2004；Miller and Hahn 2006）。TBP 结合在 TATA 盒的小沟中，并使 DNA 朝大沟弯曲 80°。该弯曲是起始位点 DNA 在裂缝中定位的必需步骤之一（图 1.10）。TFIIB 由两个结构域组成：氨基末端锌带结构域和羧基末端核心结构域，二者由一个柔韧连接分开，该柔韧连接含有一个称为 B 指的保守区。TBP/TFIIB 和 TFIIB/Pol II 的化学交联与晶体结构的结合，揭示出 TFIIB的 B 指从 RNA 出口通道进入，并结合在邻近 Pol II 催化位点的裂缝中。TFIIB 的羧基末端结构域结合 TBP 及 Pol II 的裂缝和内壁区域，这种结合会使离开 TBP 的 DNA 导向活性裂缝。在这个阶段，由于缺乏晶体结构，细节还不清楚。化学交联研究将 TFIIH 的 XPB 解螺旋酶亚基置于 TSS 下游。XBP 解螺旋酶引起处于下游位置和 TFIIB 结合位点间的 DNA 上产生负超螺旋应变，从而导致在起始位点上 11～15bp 的 DNA 熔解（Kim et al. 2000；Hahn 2004），这使得模板链插入催化裂缝中。某

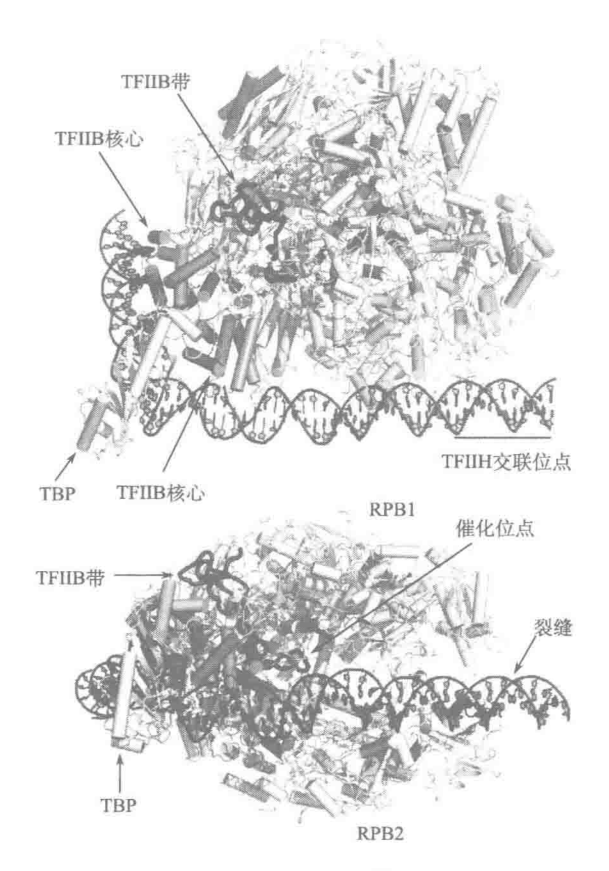


图 1.10　TFIIB 和 TBP 在邻近 Pol II 裂缝的启动子 DNA 上的定位（坐标来自 S. Hahn 及改编自 Miller and Hahn 2006）。图中展示了由 TBP、TFIIB、Pol II 及 DNA 形成的复合物的两个视图。TBP 将 DNA 向大沟弯曲并围绕在 Pol II 周围。TFIIB 核心结合邻近 TBP 和 Pol II 的 DNA。TBP 和 TFIIB 将 DNA 导向 Pol II 裂缝。（粗深灰色条带）TFIIB 氨基酸末端结构域的位置。（上面的视图）由 TBP 产生的弯曲和 TFIIB 核心的两个结构域的定位；（下面的视图）TFIIB 氨基末端的结合及其靠近催化位点的定位（由 J. Heiss 绘制，UCLA）。

些证据表明 TFIIF 和 TFIIE 在非模板链紧靠聚合酶的定位中发挥作用，以形成被称为开放复合物的稳定转录泡（Kornberg 2007）。研究表明，在 TSS 上带有预先形成的转录泡的模板避开了对 TFIIH 和 TFIIE 的需要，但仍需要 TFIIF（Holstege et al. 1995）。

第一个磷酸二酯键形成后，DNA 在桥梁螺旋的弯曲的驱动下开始移位，并且顺序通过活性位点。对延伸 Pol II 结构中 B 指的建模表明，延伸 RNA 的第 5 个核苷酸之外，在 TFIIB 和生长的 RNA 链之间，似乎有空间上的冲突。已经推测，为了继续生产性延伸，TFIIB 需要从 PIC 上分离下来。晶体结构中的这个距离类似于已经被描述为流产性起始的距离；在流产性起始中，Pol II 合成短的转录物但没有离开或逃离启动子。

因此，TFIIB 的出口可能对 Pol II 的启动子逃脱很重要。

在酿酒酵母提取物中的相关研究已发现，一旦开始生产性延伸，TFIIB 的确连同正在延伸的 Pol II 及 TFIIF 从 PIC 释放，留下激活因子、调解因子、TFIID、TFIIE 和 TFIIH（Hahn 2004）。保留在核心启动子上的整套因子被称为再起始支架。该支架复合物帮助并加快激活因子依赖性再起始过程。Pol II 的晶体结构表明，CTD 紧挨着 RNA 出口通道的外面定位。当 RNA 从通道中出现时，对其以 7meGTP 加帽，此过程需要 TFIIH Cdk7 亚基对 Pol II CTD 在 Ser-5 上的磷酸化。该磷酸化为 mRNA 加帽机器提供结合位点。

调解因子

大多数当前的研究都支持这样一种想法，即在酿酒酵母和哺乳动物细胞中调解因子是 PIC 的主要共激活因子（Biddick and Young 2005；Conaway et al. 2005；Kornberg 2005；Malik and Roeder 2005；Roeder 2005）。调解因子复合物最早被鉴定为酿酒酵母提取物中的一种生化活性物质，该活性物质可使激活因子刺激包含纯化的 GTF 和Pol II 的基础转录反应。实际上，调解因子不仅可使转录激活，而且在缺乏激活因子的情况下，体外高水平的基础转录也需要调解因子。酿酒酵母的调解因子由 25 个亚基组成，而哺乳动物的调解因子含有大于 30 个亚基（关于全面的进化对比，见 Bourbon [2008]）。调解因子的一个或多个亚基对大多数酿酒酵母基因的转录是必需的，而且很多调解因子亚基的敲除在小鼠中是致死性的。结构和遗传研究均表明，调解因子不仅共激活，而且根据环境还共抑制转录（Ding et al. 2008），并在基因调控中可能发挥其他起始后作用（Wang et al. 2005）。

正如通过 EM 所揭示的，调解因子的结构由 4 个模块组成，分别称为中部、头部、尾部及激酶（Chadick and Asturias 2005）。酿酒酵母调解因子和哺乳动物调解因子的结构相似。最大型的哺乳动物调解因子在其最长和最宽点上约为 42 nm×18.5 nm。相比之下，Pol II 只有 14 nm×13.6 nm×11 nm。显然，调解因子可基本包裹住 Pol II。调解因子显示椭圆形，而在 Pol II 出现时，该椭圆形状则改变构象，呈新月形，从而包围更接近球形的 Pol II（图 1.11）。

有数个研究已鉴定出至少两种形式的调解因子。一种形式含有带有 CDK8 或 CDK11 的激酶模块和极少的 Med26；另一种形式含有 Med26 和接近化学当量数量的 Pol II，但几乎不含有激酶模块（Paoletti et al. 2006）。某些数据表明激酶形式是抑制性的，而 Med26/Pol II 形式是刺激性的。然而，其他研究显示激酶活性有助于体外基因活化，并可能在体内也促进基因活化（Liu et al. 2004）。Pol II 似乎凭借其 CTD 和核心酶区域与调解因子结合。另外，除了与 Pol II 相互作用，调解因子在启动子 DNA 环境中也与 TFIID 接触（Johnson et al. 2002）。

很多不同的激活子接触调解因子，并且这种相互作用可能对调解因子向核心启动子的募集很重要（Biddick and Young 2005；Chadick and Asturias 2005；Conaway et al. 2005；Kornberg 2005；Malik and Roeder 2005）。已经以哺乳动物复合物对该问题进行了更为广泛的研究（见第 10 章和第 12 章）。例如，通过在中间模块中发现的 MED1 亚

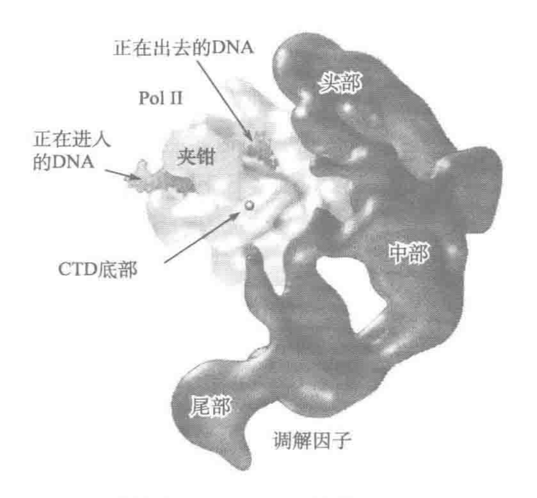


图 1.11　调解因子/Pol II 结构模型。包围 Pol II 的调解因子 Cryo-EM 结构（经 Macmillan Publishers Ltd 许可，改编自 Hahn 2004）。

基中的 LXXLL 模体，调解因子与核受体相互作用，包括甲状腺受体（TR）和维生素 D 受体（VDR）。尾部模块中的 MED15 与 SMAD2/SMAD4 及 SREBP 相互作用。高效单纯疱疹病毒 1 型激活因子 VP16 与 MED25 相互作用，而腺病毒激活因子 E1A、细胞激活因子 ELK1 与 MED23 相互作用。MED23 和 MED25 的位置还未确定，但是基于 MED23 敲除小鼠胚胎成纤维细胞（MEF）中 MED16 和 MED24 的不稳定性，以及在 MED24 敲除 MEF 中 MED23 和 MED16 的不稳定性，MED23 可能存在于尾巴模块中。因此，MED23 和 MED24 似乎通过 MED16 连接到尾巴模块上。最后，定位在头部结构域的调解因子亚基 MED29 与果蝇的兼性蛋白（intersex protein）相同，该兼性蛋白直接与序列特异性激活因子双性蛋白（doublesex）相互作用。这些观察结果共同指向这样一种想法，即除了激酶模块，所有核心调解因子模块均可参与与激活因子的相互作用。显然，凭借激活因子的调解因子募集在体外是一个限速步骤，并且有可能在体内也是如此，但这并不排除调解因子也可能在起始后步骤中起作用的想法（Struhl 2005）。

正如 EM 所测量的，调解因子的构象随不同激活因子和 Pol II 的结合显著改变（Taatjes et al. 2004a，b）。这种构象改变的机制和功能后果均不清楚，但是已经推测这些构象改变影响 Pol II 的活性或者 PIC 的活性。尽管如此，大量 ChIP 和生化研究非常清楚地表明，激活因子在体内外募集调解因子，而且这种募集增加了 GTF 的稳定性并增强了其功能。

TFIID 和 TAF

如上所述，通过结合 TSS 周围大量序列元件，TFIID 是将调解因子和剩余 GTF 拴系到核心启动子上的关键因子。通过 EM 研究，已揭示了酿酒酵母和人的 TFIID 结构。

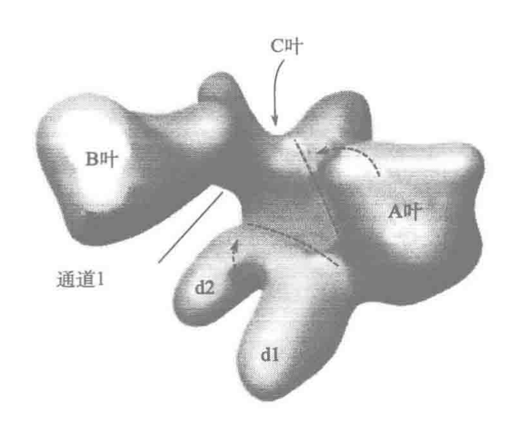


图 1.12　TFIID 的 Cryo-EM 结构。TFIID 是具有很大柔韧性及大量通道和洞穴的三叶状结构（A、B、C）。TBP 的大体位置是在 C 叶底部的主要洞穴中。TEIIB 可能结合在通道 1 附近并紧靠 B 叶。TEIIA 在 A 叶和 d1 之间（经 Elsevier 许可，改编自 Grob et al. 2006；图片由 Eva Nogales 提供）。

TFIID 是三叶马掌形结构（20 nm×13.5 nm×11 nm），其中 TBP 定位于中央空腔的底部，据认为，DNA 穿过该中央空腔（图 1.12）（Andel et al. 1999；Grob et al. 2006）。可能由于三叶间铰链的移动，该复合物承受相当大的构象柔韧性。该结构包括几个通道和洞穴，且这些通道和洞穴是其他蛋白质相互作用的表面。RNA Pol II 转录需要在 TFIID 中出现 13～14 个 TAF（Hahn 1998；Albright and Tjian 2000；Chen and Hampsey 2002；Tora 2002；Davidson et al. 2005）。注意，对于所有的三个核 RNA 聚合酶的转录，均需要 TBP，但是在 Pol II 和 Pol III 转录的环境中，一套不同的 TAF 结合到 TBP 上。TAF 在 TBP 复合物外也发挥作用。在诸如酿酒酵母的 SAGA 和哺乳动物的 STAGA/TFTC 等染色质修饰酶中，发现部分 TAF。

从酿酒酵母到哺乳动物，TAF 呈一定的保守性。虽然核心 TAF 似乎普遍存在，但也有细胞特异性 TAF 的几个例子，包括 5 个果蝇雄性生殖细胞特异的并与参与精子发生的旁系同源物。还存在几个其他旁系同源物，其中有些具有基因表达的组织限定模式，并且有很多 TAF 的拼接异构体。TAF 涉及核心启动子识别、与其他 GTF 的相互作用、激活转录，以及染色质识别。我们已在上文讨论过它们作为核心启动子识别因子的作用，并在这儿讨论它们在激活转录和染色质中的作用。

TAF 的精确作用和生理功能已是很多研究和辩论的主题。最早的提议之一是 TAF 是转录激活所必需的。的确，在哺乳动物细胞和果蝇中的生化及遗传学研究揭示，TAF 直接与激活因子相互作用，并且在体外介导这些激活因子的作用（Albright and Tjian 2000）。在少数例子中，在体内确证了对激活的这种需求。在果蝇、哺乳动物和酿酒酵母生化系统中的研究表明，激活因子将含有 TFIID 和 TFIIA 的复合物募集到核心启动子上，该效应需要 TAF（如见 Chi et al. 1995）。果蝇中的遗传学实验对该生化实验提供了有力支持，这些遗传学实验表明发育期间 TAF 在活化转录中发挥作用（Albright and Tjian 2000）。例如，激活因子 Dorsal 在体内同 TAF4 和 TAF6 相互作用；遗传学研究表明，TAF 对于 Dorsal 介导的体内蜗牛（snail）基因和扭曲（twist）基因的活化是必需的。

酿酒酵母中的其他研究也表明，TAF 是非常重要的基因，并且通过温度敏感型突变体对 TAF 的瞬时失活废除了转录。全基因组研究表明，多达 84%的酿酒酵母基因需要一个或多个 TAF。对 TAF 应答性启动子和 TAF 非依赖性启动子的分析表明，TAF 应答性基因的上游激活序列（UAS）决定 TAF 的依赖性。例如，将来自 TAF 依赖性启动子的 UAS 融合进 TAF 非依赖性基因的核心启动子区，结果表明，在这些

融合体上，UAS将TAF募集到TAF非依赖性启动子上。这些数据说明，在这些UAS上激活因子直接与TAF接触（Chen and Hampsey 2002）。而且，在酿酒酵母和哺乳动物细胞中进行了TAF的大量ChIP研究，在这些研究中，TAF的募集与基因的活化状态有关，甚至在无活性状态中单独的TBP可能被预先结合的情况下，也是如此。

9种TAF含有组蛋白折叠结构域。TAF和组蛋白间显著的同源性导致这样一种假说，即可能作为一种稳定它与DNA的结合的手段，或作为一种转录复合物组装期间置换核小体的方式，TFIID可能模拟核小体的功能。已经形成了5对带有组蛋白折叠的TAF（4-12、6-9、3-10、10-8及11-13）。不同的TAF对遍布TFIID的三个叶，其中某些TAF对呈多重拷贝（Leurent et al. 2002）。的确，TFIID/DNA复合物的化学交联和DNA拓扑分析表明，DNA可能缠绕在TFIID周围，可能像DNA缠绕在一个八聚体周围那样（Hoffmann et al. 1996；Oelgeschlager et al. 1996）。然而，对核小体晶体结构内与DNA相互作用的组蛋白表面的检查表明，介导关键接触的残基（主要是与小沟及磷酸骨架在一起的精氨酸）在TAF中并不是很保守。尽管有这样的分析，仍有报告显示各种组蛋白折叠TAF结合DNA（即TAF6-9和TAF4b-12），从而表明TAF使用其他一些替代表面识别核酸（Shao et al. 2005）。另一种看法是TAF作为一种使TFIID与附近核小体对接的手段，使核小体间发生接触。

除了TAF模仿八聚体结构的可能远程遥控外，某些TAF具有识别被修饰染色质的结构域。例如，TAF1含有一个可在体外结合乙酰化H4尾巴的一双溴域（Jacobson et al. 2000）。然而，要注意，某些溴域也识别转录因子内的乙酰化残基。因此，TAF1也结合乙酰化p53，从而提供了一种TFIID向启动子的募集机制（Li et al. 2007b）。同样，TAF3含有结合H3K4me3的PHD（植物同源域）指，从而提供了TFIID和MLL/SET1复合物间的纽带，MLL/SET1复合物在基因活化期间被招募到启动子上（Vermeulen et al. 2007）（见下文）。

综上所述，TAF在转录中有很多作用。它们不仅对启动子识别是必需的，而且结合特定的染色质修饰，以及与某些激活因子直接相互作用。另外，已报道TAF1含有HAT和激酶活性，认为这对基因活化很重要。

活化和抑制

基因活化

应当强调的是基因可在不同的步骤被调控，这些步骤包括起始、延伸、mRNA加工和稳固、向细胞质的运输及翻译。我们在此聚焦于起始和早期延伸水平上的基因活化过程。被称为激活因子的蛋白质刺激所有mRNA编码基因。激活因子一般是小的模块化的蛋白质，这些蛋白质拴到近端启动子和增强子的特异位点上（Ptashne and Gann 1990）。因为基因活化从激活因子开始，所以在20世纪80年代和90年代很多研究的重点都放置在结构、模块设计和激活因子的调控上。大多数激活因子含有可分离的DNA结合及活化结构域，其中DNA结合结构域负责启动子/增强子的结合（见第10章和第

11 章)，而激活结构域接触染色质修饰/重塑蛋白质及共激活因子（见第 10 章、第 11 章及第 12 章）。所有受调控基因都对以激活因子开始的信号或者指示作出应答 (Ptashne and Gann 1990)。

激活因子在数种不同水平上被调控。某些激活因子直到它们接收到特定信号时才被合成、活化或定位到细胞核。例如，糖皮质激素受体（GR）结合一个小的类固醇配体，该配体导致 GR 在细胞质中与伴侣分子分离并定位到细胞核中（Lefstin and Yamamoto 1998）。其他诸如 RARRXR 异源二聚体的核受体仍然以抑制形式结合到应答启动子上并被配体激活（Xu et al. 1990）。另一种激活 DNA 结合激活因子的方式是通过翻译后修饰。Elk-1 反式激活因子结合 DNA，但直到被有丝分裂原活化蛋白激酶（MAPK）磷酸化时才被激活（Treisman 1996）。某些激活因子是细胞和发育特异的。无论情况如何，可能由于它们的小尺寸，通过结合核小体间的连接区或者核小体表面暴露的 DNA 的短的伸展，这些 DNA 结合蛋白确实显得很少接近染色质。同时，已知在某些情况下，激活因子协同性地结合到启动子和增强子上，以形成稳定的复合物，这些复合物称为增强体（在第 11 章中评述）。一旦激活并结合，这些蛋白质则起始一系列导致基因活化的事件。

虽然该过程好像十分简单，但是我们对任何一个基因被激活的精确机制只有有限的了解。在酿酒酵母中有一些研究得最彻底的例子，酵母中启动子-增强子结构更简单，并且遗传学和 ChIP 的使用能确证生化研究结果。从这些研究弄清楚的是，在募集 Pol II机器前，激活因子必须首先改变增强子和启动子处的染色质（如见 Reinke and Horz 2003）。酿酒酵母中启动子区好像有更低的核小体密度，并且在基因活化期间核小体很明显被驱走（见下面）（Li et al. 2007b）。

激活因子介导的 Pol II 机器募集到核心启动子上后，通过与 TFIID 及调解因子的相互作用，Pol II 必定会在染色质模板上起始和延伸转录（Li et al. 2007a）。核小体充当 Pol II 延伸的固有障碍，因此，Pol II 必须指导转录期间阻碍性核小体的重塑。重要的是，核小体在转录后必须返回到它们最初的位置，以防止来自基因体内的伪转录起始。最后，转录过程中，Pol II 必须同 RNA 加工机器联系。因此，已出现的详细资料表明该过程是一系列非常复杂的催化事件，这些事件涉及很多类型的蛋白质修饰。

转录起始期间的染色质修饰和重塑

在很多情况下，启动子的乙酰化和重塑早于转录（Workman 2006）。正如上文所讨论的，HAT 介导乙酰化，而 ATP 依赖性酶介导重塑，下文将更详细地讨论这些论题。在启动子处于积极沉默状态（即被 Polycomb 或 HP1 结合，见下文）的情况下，去甲基化和抑制标记的去除与组蛋白超乙酰化和 H3K4 三甲基化同时发生。

通常，基因活化的最初步骤之一是激活因子将各种 HAT 募集到基因上（图 1.13）。这些 HAT 包括含有 GCN5（即哺乳动物中的 STAGA 和 TFTC 及酿酒酵母中的 SAGA 及其变体）或 p300 及 CBP 的复合物等（Lee and Workman 2007）。例如，β-干扰素（IFN-β）增强体结合 p300 并将其募集到启动子上（Merika and Thanos 2001）。的确，

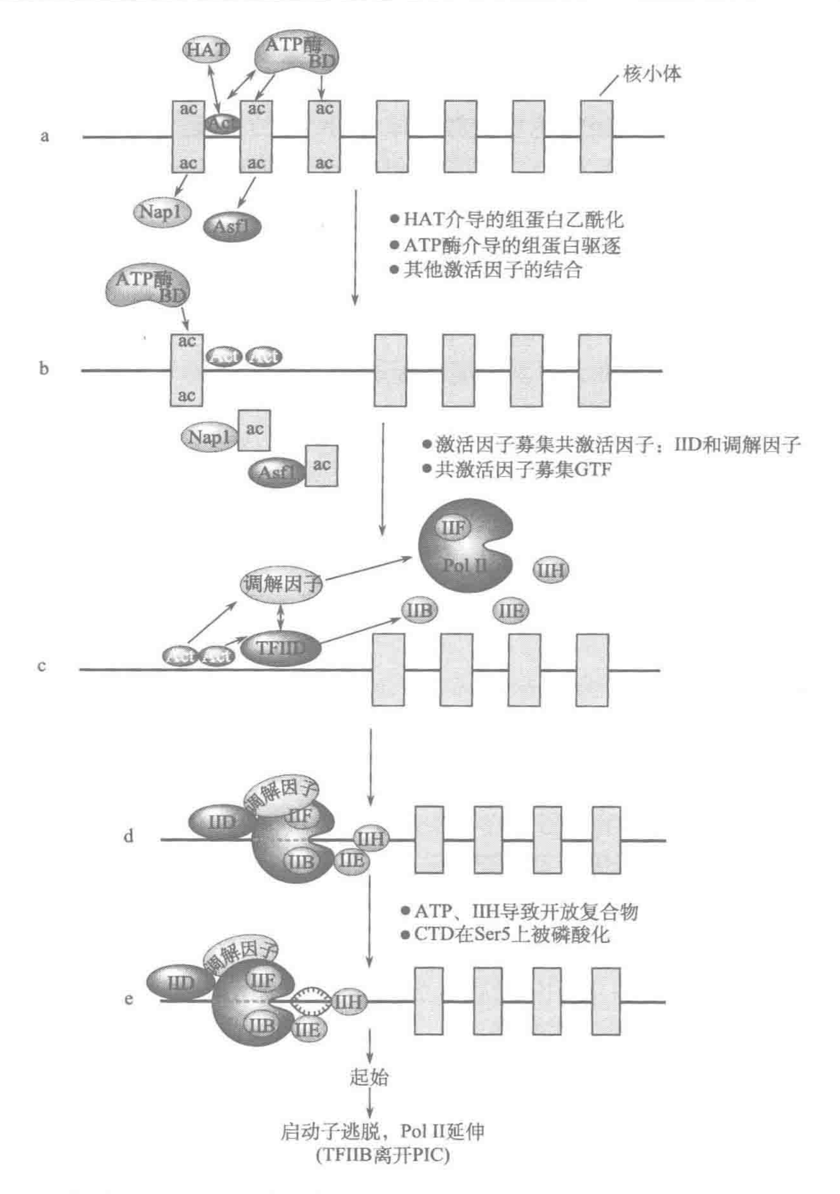


图 1.13 基因活化。基因活化中的各种步骤如图所示。(a) 激活因子（Act）结合到近端启动子中的DNA位点上，并且募集ATP依赖性重塑酶（ATP酶）和组蛋白乙酰转移酶（HAT）。HAT乙酰化核小体上的组蛋白尾巴（ac）。ATP酶也可通过其溴域（BD）结合乙酰化的赖氨酸。(b) 诸如Nap1和Asf1的伴侣分子充当被ATP酶驱逐的组蛋白的受体。其他的激活因子可能在组蛋白移除后结合。(c) 激活因子之后募集像调解因子和TFIID这样的共激活因子。TFIID在TFIIA的帮助下结合TATA盒（TFIIA没有显示）。(d) 共激活因子募集通用转录因子和Pol II。TFIIH（IIH）结合起始位点的下游。(e) 由IIH引起的ATP水解导致起始位点的熔解，以形成开放复合物。IIH的Cdk7亚基在Ser5上磷酸化Pol II CTD。Pol II起始转录，并且TFIIB与正在延伸的Pol II和TFIIF一起离开复合物。当Pol II向基因内延伸8～10个核苷酸时，它从启动子上逃脱。其他细节在正文中讨论。请注意受增强子影响的步骤还不是很清楚（本图未按比例绘制）。

p300 的结合似乎在体外稳定增强体。一旦结合，HAT 就会乙酰化组蛋白尾巴及其他蛋白。

被乙酰化的尾巴破坏高级染色质接触（即 H4K16ac）的稳定性，并充当其他含溴域蛋白的识别位点（Taverna et al. 2007），包括 ATP 依赖性重塑蛋白（Peterson and Workman 2000；Mohrmann and Verrijzer 2005；Li et al. 2007a）。例如，酿酒酵母中 SWI/SNF 复合物含有识别乙酰化赖氨酸的溴域。酿酒酵母的 RSC 染色质重塑体含有 8 个识别各种乙酰化组蛋白残基的溴域。在哺乳动物中，PBAF 含有至少 7 个溴域，这与在 RSC 亚基中发现的相似（Mohrmann and Verrijzer 2005）。正如上文所描述的，RSC 和 SWI/SNF 通过很强的亲和力识别乙酰化赖氨酸，并且这两个酶均可重塑并从模板上驱逐核小体。很多研究说明核小体在基因活化期间从体内酿酒酵母的启动子上被驱除，这也是通用机器募集的先决条件（Workman 2006）。虽然上文讨论的模型暗示 HAT 比 ATP 依赖性重塑机器出现得早，但也有发生相反的情况（Cosma 2002）。激活因子在酿酒酵母和哺乳动物细胞中可直接接触 SWI/SNF，而且可信的是，重塑蛋白的活动对于松弛高级结构染色质以允许乙酰化是必需的。有可能乙酰化的组合及与激活因子的直接接触将重塑酶募集到启动子上。最终的目标是为通用转录机器的募集生成一个平台。

通用机器募集的一种模式

在综述文献的时候，常见共激活因子这个术语非常广泛地用于指示任何与序列特异性 DNA 结合蛋白直接结合的物质。在某些情况下，这些蛋白质涉及染色质重塑，而在其他情况下，所谓的共激活因子可能仅仅是实际的活化结构域。在本章中，我们将共激活因子这一术语限定为调解因子及 TFIID 复合物，因为它们的结合对于转录起始是限速的，并且，正如上文所讨论的，已知 TAF 和调解因子亚基是激活因子的直接靶标。在生化研究中，调解因子和 TFIID 同启动子 DNA 上的激活因子形成复合物，并募集剩余的 GTF。这些募集事件并不需要 ATP，并且一旦剩余的 GTF 结合，前起始复合物同未熔解的 DNA 呈关闭形式。ATP 的加入为数个关键事件发出信号，这些事件使 Pol II形成开放复合物并起始转录；在 Pol II 开放复合物中，起始位点熔解。正如上面讨论的，DNA 熔解需要通过 TFIIH XPB 亚基水解 ATP。有一些激活因子可能影响该熔解步骤的证据。

在起始期间或起始后不久，调解因子的 CDK8 亚基和 TFIIH 的 CDK7 亚基磷酸化 Pol II CTD 的 7 残基重复（heptad repeat）内的丝氨酸（Phatnani and Greenleaf 2006）。CDK7 磷酸化 Ser-5 需要两个事件：mRNA 加帽机器的募集及 PAF1 复合物的募集，PAF1 复合物依次对 SET1/MLL 复合物的招募是必需的，而 SET1/MLL 复合物三甲基化 H3K4（Shilatifard 2008）。

在此结合点上，一系列难以理解的观察结果似乎与上文描述的模式相矛盾。第一个警告是哺乳动物 TFIID 有串联溴域和 PHD 结构域，二者均可结合被修饰的染色质。因此，如果这些模块在基因活化期间确有功能，那么这将暗示核小体仍然出现在启动子上，这与移除核小体为通用机器让路的简单看法相矛盾。第二个警告是通用机器必须结

合在启动子处以磷酸化 Pol II CTD，这对三甲基化 H3K4 的 Set1/MLL 复合物的结合是必需的。这似乎将暗示 H3K4 三甲基化发生在 TFIID 结合之后。因此，H3K4me3 协助 TFIID 招募的想法似乎不可能，相反，却暗示这种相互作用具有起始后作用（即结合的稳定维持）。

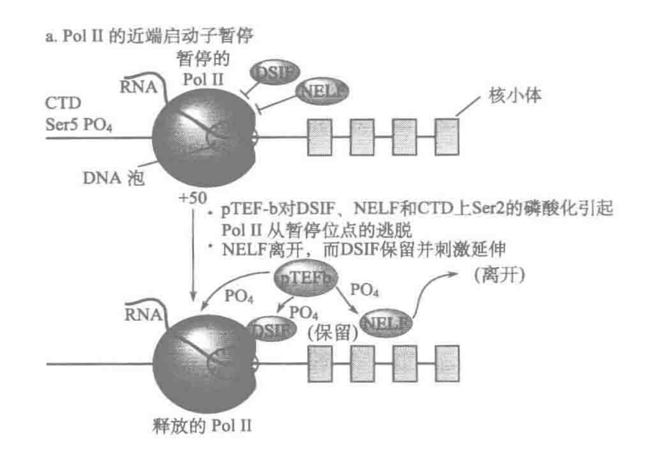
值得注意的是，在酵母中一个被称为 Bdf1 的单独的蛋白质编码在哺乳动物 TAF1 中发现的溴域（Matangkasombut et al. 2000；Taverna et al. 2007）。更重要的是，在酵母的 TAF3 中没有 PHD 结构域。因此，在酵母和哺乳动物中，各种复合物的机制和功能顺序可能不同。不像 TAF1 和 TAF3，CHD1 从酵母到人是保守的（Daniel et al. 2005；Hall and Georgel 2007）。正如上文所描述的，CHD1 是 SWI/SNF 家族的含染色域的 ATP 酶。CHD1 通过其染色域也结合 H3K4me3。H3K4 的三甲基化同活性基因的 5′ 端有联系，认为它在延伸中发挥某些早期作用。

对上面募集模式的第二个警告是某些哺乳动物启动子似乎被一个或多个 GTF 和 TBP 提前结合（Cosma 2002）。最初的事例涉及在肝细胞分化的漫长过程中被激活的基因，这些基因包括抗胰岛素 α1（Soutoglou and Talianidis 2002）和肝细胞核因子-4α（Hatzis and Talianidis 2002）。在这些情况下，部分激活因子和 GTF 已出现在启动子上，并且新激活因子的出现导致 TAF、调解因子复合物、全套 GTF、染色质修饰复合物和 Pol II 的募集。

聚合酶 II 延伸的初始阶段

Pol II 高效延伸的主要障碍之一是暂停（Shilatifard et al. 2003；Peterlin and Price 2006；Saunders et al. 2006；Armstrong 2007；Svejstrup 2007a）。在起始子和启动子逃脱后，Pol II 延伸到相对 TSS 的＋20 和＋50 之间并暂停（图 1.14）。该暂停依赖于被称为 DRB 敏感性诱导因子（DSIF）和负延伸因子（NELF）的两种因子。该暂停似乎是延伸过程中的一个关键检测点，可能使 Pol II 从结合 mRNA 加帽机器转换为结合拼接及转运机器。延伸因子 pTEF-b（酿酒酵母中为 Ctk1）的 CDK9 亚基对 Pol II 从该暂停上的释放是必需的（Peterlin and Price 2006）。数条证据表明激活因子募集 pTEF-b 并影响延伸的这一早期阶段。一旦募集，CDK9 磷酸化 DSIF（酿酒酵母中为 Spt4 和 Spt5）、NELF 及 CTD 的 Ser2。磷酸化的 NELF 离开，但是磷酸化的 DSIF 保留并刺激延伸。已经表明在酵母中 DSIF 的 Spt4 亚基有助于 PAF1 复合物结合 Pol II（Qiu et al. 2006）。这种观察结果暗示上文描述的 PAF1 依赖性 H3K4 三甲基化事件可能与 Pol II 延伸的初始阶段关系更密切，这种关系的密切程度超过了这些事件与转录起始的关系。

学术界对 Pol II 正好暂停在起始位点下游一直有相当大的兴趣（Margaritis and Holstege 2008），很多无活性基因也含有暂停的 Pol II。已提出这种暂停 Pol II 起维持基因下游染色质开放以便快速再激活的作用，另外，一旦基因激活，暂停的 Pol II 执行第一轮或先驱轮的转录。同时还提出这种 Pol II 仅仅是在所有基因上发生的一个偶然的和随机的事件，但发生这种事件的基因只占很小的比例。两种观点都有支持者，但有趣的是，ChIP 实验的观察结果表明更多的处在暂停位点上的 Pol II 在活性基因上，而不是在无活性基因上，如果在 100％的细胞中没活性的基因都被暂停的 Pol II 结合，这将



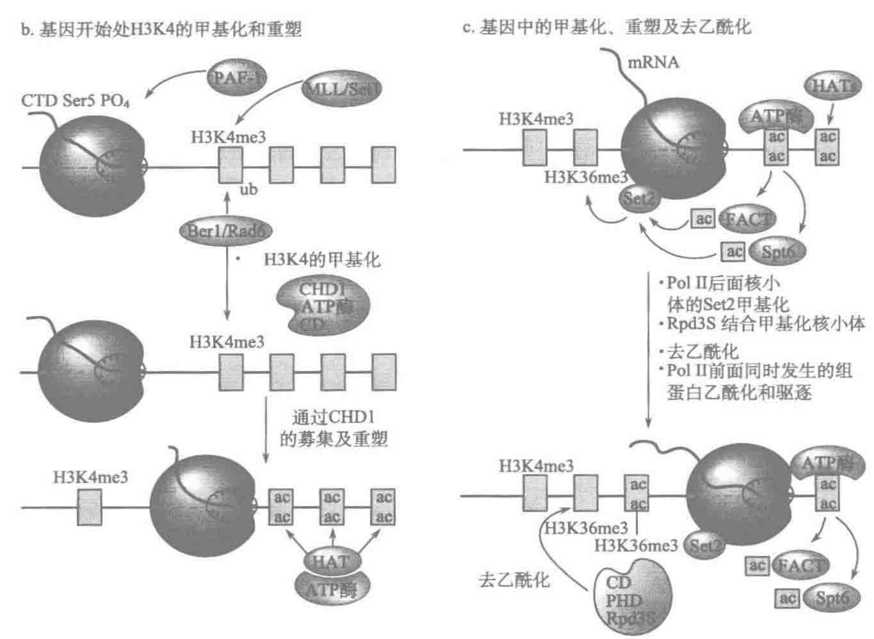


图 1.14 基因上的启动子逃脱、暂停及延伸。Pol II 延伸中的各种步骤如图所示。(a) Pol II 从启动子上逃脱后，暂停在转录起始位点下游（+50）大约 50 bp 处。暂停是由被称为 DSIF 和 NELF 的两种因子的作用引起的。延伸因子 pTEF-b 磷酸化（PO_4）DSIF、NELF 及 CTD 的 Ser2。NELF 离开暂停位点，而 DSIF 保留并刺激延伸远离暂停位点。(b) 在延伸的初始阶段，可能同时发生或先于 a 中所描述的步骤，PAF-1 复合物被募集到 Pol II 上。PAF-1 募集 MLL/Set1 复合物（酵母中为 COMPASS），该复合物三甲基化 H3K4。该甲基化（H3K4me3）取决于由 Brel 和 Rad6（在哺乳动物中为 RNF20 和 RNF40）引起的 H2B 泛素化（Ub）。H3K4me3 通过其染色域（CD）募集 CHD1 ATP 酶。CHD1 重塑转录起始位点附近的核小体。HAT 乙酰化下游核小体，并且 ATP 酶重塑/驱除核小体。关于该步骤如何发生知之甚少。(c) 延伸期间，诸如 Spt6 和 FACT 的伴侣在 Pol II 后面组装核小体。Pol II 延伸需要组蛋白乙酰化，但是细胞已发展了一种系统，用于乙酰化拖尾核小体，以终止 Pol II 通过后的重塑。在该途径中，Pol II CTD 在 Ser2 上发生磷酸化（PO_4），这将引起 Set2 的结合。Set2 在 H3K36（H3K36me2）上三甲基化拖尾核小体，这种修饰为 Rpd3S 组蛋白去乙酰化复合物产生一个结合位点，该复合物通过它的染色域（CD）和 PHD 结构域结合 H3K36me3。之后 Rpd3 去乙酰化亚基去除乙酰基团，并阻止 ATP 酶重塑拖尾核小体。拖尾核小体去乙酰化阻止隐秘起始。其他细节在正文中进行了讨论（本图未按比例绘制）。

没有意义。无论哪种方式，该问题是有趣的，并且是很多研究的主题。

从近端暂停位点逃脱后，Pol II 延伸进基因内。Pol II 似乎在转录期间频繁地暂停，暂停的原因不完全清楚，但可能归因于某些 DNA 序列、结合蛋白的出现，或核小体的出现。然而，在暂停期间，RNA 链的 3′ 端可能会与模板错位。很多不同蛋白质因子在固有的暂停位点处起作用，或阻止暂停，或逆转暂停（Shilatifard et al. 2003；Svejstrup 2007a），这些因子包括 SII（TFIIS)、TFIIF、Elongin 及 ELL（Shilatifard et al. 2003）。

据认为，一旦暂停，Pol II 则原路返回，并且错位的 RNA 3′端从孔中被逐出，故认为该孔就是核苷酸进入的孔（Svejstrup 2007a)。通过刺激 Pol II 的转录本切割活性，TFIIS 积极逆转暂停 Pol II/TFIIS 的结构已经清楚（Cramer et al. 2008）。TFIIS 结合在 Pol II 孔中，并延伸到激活位点。据认为，TFIIS 重排活性位点并放置一个金属离子加一个水分子，以刺激错位的 RNA 3′端水解。该事件使 Pol II 从新的正确对齐的 RNA 3′端继续转录。如果不能通过暂停，则认为 Pol II 被泛素化并降解（Svejstrup 2007a)。ELL 和 Elongin 极大地刺激体外延伸的速度，并且认为两者可能阻止导致暂停的 RNA 错位。在酿酒酵母中发现并命名为延蛋白（elongator）的另一个复合物也刺激延伸(Svejstrup 2007b)。延蛋白有 HAT 活性，这对染色质上的延伸可能很重要。在细胞质中也发现了延蛋白，在细胞质中它发挥数种不同的作用。

聚合酶 II 在基因内遭遇核小体

当 Pol II 进一步延伸进基因时，会遭遇核小体。很多研究已表明核小体在基因体中的密度比启动子中高。然而，在已进行了全基因组 ChIP 研究的酿酒酵母中，转录活性和基因内核小体密度间似乎呈负相关（Workman 2006)。这一观察结果表明在转录延伸期间，基因内发生重塑和某些驱逐（Armstrong 2007；Li et al. 2007a)。通过脉冲标记或掺入新合成的表位标记（见第 10 章）的组蛋白，测定组蛋白的周转率，结果表明在转录的基因中 H2A/H2B 周转水平非常高，而 H3/H4 周转水平较低但显著。在哺乳动物细胞中，H3.1 是占主导地位的 H3 种类，但在转录期间 H3.1 被 H3.3 取代，从而表明 H3 周转确实发生（Armstrong 2007)。其他研究表明 H3.3 周转速率比 H3.1 更高。热激基因的研究巩固了核小体在转录期间被解开的想法。在酿酒酵母和高等真核生物中，热激基因显示了横跨整个基因座的核小体显著缺失，即使在 Pol II 已开始转录进基因前也是如此。

组蛋白取代的机制还不清楚，但是活性基因在编码区被乙酰化，并且 ATP 依赖性重塑酶是转录必需的（Saunders et al. 2006；Workman 2006)。另外，一组不同的组蛋白伴侣对延伸期间组蛋白的移除和取代是必需的。酿酒酵母中组蛋白伴侣的突变体对延伸有特定的影响，这些突变体包括 Asf1、Spt6 和 FACT（酿酒酵母中是 Spt16 和 Pob3)（Saunders et al. 2006；Workman 2006；Armstrong 2007)。对组蛋白伴侣的需要与一种事实相关，即 ATP 依赖性重塑蛋白在受体缺乏时不从 DNA 上驱走组蛋白，这归因于组蛋白八聚体同 DNA 的高度亲和力。在生化试验中，受体可包括其他 DNA 分子或组蛋白伴侣。例如，Nap1 可接受被 RSC 驱走的组蛋白（Lorch et al. 2006)。在

Nap1 或受体 DNA 缺乏时，RSC 将重塑核小体（即将其滑动到新位置），但不驱逐八聚体。因此，在体内，必须有伴侣来帮助移走或驱逐八聚体、H3/H4 四聚体或 H2A/H2B 二聚体。

已报道，在哺乳动物生化系统中，FACT 在体外解开核小体以促进延伸（Reinberg and Sims 2006）。然而，酿酒酵母 FACT 在体外是一个非常强有力的核小体组装因子。同样，酿酒酵母 Spt6 在体外从游离组蛋白组装核小体，但好像不驱逐八聚体（Bortvin and Winston 1996）。然而，FACT 随同 NHP6A 的确显示在没有 ATP 依赖性重塑蛋白的情况下显著重塑核小体（Rhoades et al. 2004）。另外，Nap1 伴侣可沿着 DNA 滑动组蛋白，并可能有助于组蛋白驱逐（Park and Luger 2006）。

Spt6 和 FACT 可能是染色质组装因子的想法受到一种观察结果的支持，即在酿酒酵母中，编码这些蛋白质的基因中的突变引起一种已知为隐秘或内部转录起始的现象（Kaplan et al. 2003）。这些相同的突变好像不严重抑制延伸。对这种观察结果的一种解释是，在 Pol II 通过后，FACT 和 Spt6 重新组装核小体，并且这些因子中的突变阻止重组装，因而使随机转录在基因体中发生。这种隐秘转录似乎依赖于 Pol II，有限的证据表明这归因于 PIC 的形成。尚不清楚 PIC 如何组装或该过程是否需要激活因子。

Pol II 延伸期间，一个复杂的途径在运转，以确保转录期间核小体的有序重组。除了上文引用的伴侣分子，该机器包括被称为 Set2 的 H3K36 三甲基化酶和被称为 Rpd3S 的组蛋白去乙酰化酶（Li et al. 2007a），其中，S 代表小型；被称为 Rpd3L 的大型 Rpd3 是抑制期间在基因的启动子上发现的，这一点将在下文讨论。Set2 酶结合 Pol II 的磷酸化 CTD，并在转录期间一起移动。

虽然未经证实，但目前的模式是当 Pol II 转录过一个区域时，伴侣在其后重新组装核小体。该八聚体的来源目前尚不清楚，可能是最初占据该位置的八聚体或者是不同的八聚体。不管哪种来源，八聚体可能被乙酰化或者遭受乙酰化，因为 Rpd3S 是去乙酰化转移酶，并且 Rpd3S 途径的突变导致转录期间编码区的超乙酰化。据认为，Set2 在 Pol II 后面三甲基化 H3K36，然后该修饰募集 Rpd3S 复合物，该募集作用是通过该复合物的染色域和 PHD 指的组合发生的，其中染色域和 PHD 指分别是由该复合物的 Eaf3 和 Rco1 亚基编码的（Li et al. 2007c）。然后，Rpd3S 在 Pol II 后面使组蛋白去乙酰化，从而有效地阻碍含溴域的 ATP 依赖性重塑酶的募集及其活动。

一个关键的问题是，为什么核小体在 Pol II 通过后需要快速重组。对此可能有数种原因，包括保护基因组 DNA 免受损害，以及阻断基因编码区内 PIC 的组装。现已知非编码 RNA 在转录本降解、翻译控制及异染色质形成等过程中发挥主要作用。因此，阻止伪转录可能对基因组调控很重要。

有人预测具有高密度转录 Pol II 的高度转录基因可能很少依赖 Set2/Rpd3S 途径，因为 Pol II 的占据实际上将代替核小体并保护基因（Li et al. 2007d）。的确，很多证据表明这种想法是正确的，该结果表明高度转录的基因同样不依赖 Set2/Rpd3S 机器。

从酿酒酵母到人，H3K36 甲基化发生在基因内，但在哺乳动物中，在活跃转录基因的编码区存在另一种极不寻常的标记。这些编码区同时含有 H3K9me2 和 H3K9me3，

并且似乎结合异染色质蛋白 1（HP1）家族的 γ 成员（Eissenberg and Shilatifard 2006）（见下文）。HP1γ 倾向于遍及常染色质，而 HP1α 和 HP1β 一般集中于异染色质中。因此，已表明 HP1γ 发挥类似于 H3K36 的三甲基化功能。另一种与活性染色质相关的修饰是 H3K79 甲基化（Shilatifard 2006），该标记由 DOT1 甲基化酶催化。在酿酒酵母中，似乎需要相同的甲基酶，以阻止异染色质扩散进转录活性区域。

总之，在 Pol II 延伸期间，核小体必须被修饰、重塑、驱逐到某种程度，并重新组装。虽然对涉及的机器知之甚多，但是精确机制却知之甚少。这一主题仍然是该领域中研究的活跃领域。

转录的沉默或抑制

虽然活化信号的去除可能是基因失活的最初步骤，但是也存在积极去除基因活化所需的修饰的机制。通常，基因沉默或抑制与五类系统相关：组蛋白去乙酰化、H3K9 双甲基化和三甲基化、H3K27 三甲基化、H4K20 三甲基化，以及 GpG 残基上的 DNA 甲基化。每一类系统都导致特定的后果，这些后果与共抑制因子或沉默蛋白的结合，以及染色质组装成与调控 Pol II 转录不相容的结构有关。

组蛋白去乙酰化

组蛋白去乙酰化控制的抑制的概念最初来源于一种观察到的现象，即相对于非转录基因，活跃转录基因的启动子及有限范围内的编码区被高度乙酰化。因此才出现了该想法，即组蛋白乙酰转移酶是基因活化所必需的，而组蛋白去乙酰化酶则参与抑制。实际上，除组蛋白外，组蛋白去乙酰化酶可使很多蛋白质去乙酰化。在某些情况下，去乙酰化导致转录活化，而在其他情况下则导致沉默（Smith 2008）。另外，我们知道，某些组蛋白去乙酰化发生在基因体内，并且是活跃转录所必需的（Wang et al. 2002）。最后，不是所有特定赖氨酸的乙酰化都与转录呈正相关。介绍了相关的警告后，下面对组蛋白去乙酰化做一个一般性的概述，因为组蛋白去乙酰化与转录抑制相关。

哺乳动物和酿酒酵母中有 4 种类型的 HDAC（Yang and Seto 2008），它们是基于主要的酿酒酵母去乙酰化酶 RPD3、HDA1 和 SIR2 分类的。类型 1 在哺乳动物中包括与 Rpd3 有关的 HDAC1、HDAC2、HDAC3 及 HDAC8；类型 2 包括与 HDA1 有关的 HDAC4、HDAC5、HDAC6、HDAC7、HDAC9 及 HDAC10。HDAC11 显示两者的特征并标记为类型 4；类型 3 包括各种 sirtuin（乙酰化酶），它们是酿酒酵母参与异染色质形成的 Sir2 蛋白的亲属。在哺乳动物中，这些 sirtuin 还包括 SIRT1-7。所有的 sirtuin 都需要 NAD 作为去乙酰化的辅因子。

大多数 HDAC 作为多亚基复合物的组分而存在，多亚基复合物通过与 HDAC 复合物的一个亚基的相互作用被募集到基因上（表 1.4）。在某些情况下，DNA 结合蛋白同该复合物共纯化，并在其他情况下，它们不共纯化。同样，HDAC 在各种不同组蛋白上的活性在很多情况下似乎依赖于一个提供特定组蛋白靶向功能的亚基。

表 1.4 HDAC 复合物及其亚基

复合物	哺乳动物	酿酒酵母		蛋白质结构域
		Rpd3L	Rpd3S	
Sin3	HDAC1、HDAC2	Rpd3	Rpd3	类型 1 去乙酰化酶
	RbAp46、RbAp48	Ume1	Ume1	WD40 重复
	Sin3A	Sin3	Sin3	PAH 模体
	Sds3、BRMS1	Sds3		卷曲螺旋
	RBP1			
	SAP30	Sap30		
	SAP18			泛素折叠
	ING1/2	Pho23		PHD 指
			Rco1	PHD 指
			Eaf3	染色域
Mi-2/NuRD	HDAC1、HDAC2			类型 1 去乙酰化酶
	RbAp46、RbAP48			WD40 重复
	Mi-2α/β			解旋酶
	MTA1-3			SANT 结构域
	MBD2、MBD3			甲基 CpG 结合
	p66α/β			
CoREST	HDAC1、HDAC2			类型 1 去乙酰化酶
	CoREST			SANT 结构域
	LSD1			SWIRM、K-去甲基化酶
	BHC80			PHD 指
	CtBP1			脱氢酶
N-CoR/SMRT	HDAC3			类型 1 去乙酰化酶
	N-CoR/SMRT			SANT 结构域
	TBL1/TBLR1			WD40 重复
	GPS2			
	JMJD2A			jumonji 去甲基化酶
	Kaiso			甲基 CpG 结合

经 Macmillan 出版社允许，改编自 Yang and Seto (2008)。

一个表征得很好的 HDAC 复合物的例子是 Rpd3 复合物，该复合物包含一些关键亚基，这些亚基在从酿酒酵母到人的不同生物中在很大程度上是保守的。在酿酒酵母中，这个复合物以大小两型存在。大型 Rpd3 复合物（Rpd3L）由 Rpd3、Sin3、Ume1、Pho23p、Sap30、Sds3、Cti6、Rxt2、Rxt3 及 Dep1 组成。酿酒酵母中很多序列特异性抑制因子可募集该复合物。例如，序列特异性抑制因子 Ume6 和 Ash1 直接招募 Rpd3L 到靶基因。正如上面所讨论的，Rpd3 也存在于涉及 Set2 通路的小型 Rpd3 复合物（Rpd3S）中。在这种情况下，该复合物包括 Rpd3、Sin3、Ume1、Rco1 及 Eaf3。正如上面所讨论的，Rpd3S 复合物是 H3K36me6 的靶点。

各种不同的序列特异性 DNA 结合蛋白同 Rpd3 复合物的相互作用往往由 Sin3 亚基介导，Sin3 亚基在酿酒酵母和哺乳动物中是保守的（Silverstein and Ekwall 2005）。在

哺乳动物中，有两个称为 Sin3a 和 Sin3b 的相关家族成员。DNA 结合蛋白或共抑制因子同 Sin3 的大多数相互作用要经过 4 个保守的配对两性螺旋（PAH）结构域，或者一个称为组蛋白去乙酰化酶相互作用结构域的结构域。

在哺乳动物细胞中，序列特异性抑制因子的 Mad 家族（第 10 章中讨论）含有一个被称为 Sin3 相互作用结构域（SID）的结构域，该结构域直接与 PAH2 接触并将含有 HDAC1 和 HDAC2 的 Sin3 复合物募集到启动子上。酿酒酵母的 Ume6 也有 SID 结构域。Ume6 是参与减数分裂的基因的抑制因子和激活因子。在缺乏共激活因子 Ime1 的情况下，Ume6 结合 Sin3 并将 Rpd3 募集到去乙酰化的组蛋白上，因此失活基因（Silverstein and Ekwall 2005）。相反，在 Ime1 出现且 Ume6 被 Rim11 激酶磷酸化时，Ume6 激活转录。这类 DNA 结合因子在抑制因子和激活因子间的转换是转录中的一个常见旋律。

另一个例子是哺乳动物核受体家族。通过被称为共抑制因子的中间调解因子，核受体（NR）与含有 HDAC3 的复合物相互作用（Silverstein and Ekwall 2005）。在没有配体的状态下，很多 NR 与称为 N-CoR 的共抑制因子或者与其高度相关的称为 SMRT（维甲酸和甲状腺激素受体的沉默中介物）的蛋白质相互作用（见第 10 章）。N-CoR 和 SMRT 处于 HDAC3 和几个其他蛋白质的一个复合物中。在缺乏配体的情况下，NR 与其 DNA 上的靶点结合，并且配体未结合的配体结合结构域（LBD）与 HDAC3 一起结合 N-CoR 或 SMRT。一旦配体结合，LBD 就改变构象，且共抑制因子会不再结合。反之，LBD 处于一种可结合 HAT 和调解因子的构象；通过 HAT 和调解因子中含有的 LXXLL 模体，这种构象可结合诸如 p300/CBP 和 SRC 家族成员的 HAT 或者调解因子（见第 10 章）。

类型 1 的 HDAC 在很多生物体中形成许多参与抑制的其他复合物，这些复合物已得到很好的表征。Mi-2/NuRD 复合物包括 HDAC1 和 HDAC2、视网膜母细胞瘤相关蛋白 RbAP46 和 RbAP48、Mi-2 ATP 依赖性重塑亚基（也称为 CHD3/4，见上文）、识别甲基化 CpG 的亚基（MBD2），以及两个结合组蛋白的其他亚基（Yang and Seto 2008）。NuRD 在很多系统中与抑制有关。CoREST 是含有 HDAC1 和 HDAC2 的另一种复合物（Lakowski et al. 2006）。CoREST 与 REST-1 转录抑制因子结合，REST-1 是一个序列特异性结合蛋白，参与关闭与神经元分化有关的基因。CoREST 也含有可使 H3K4me2 和 H3K4me1 去甲基化的胺氧化酶 LSD1。

虽然我们已经关注了类型 1 的 HDAC，但是类型 2 的 HDAC 也形成多亚基复合物，这些复合物是 DNA 结合蛋白的靶标（Yang and Seto 2008）。某些这些复合物是肌细胞增强子因子 2（MEF2）的靶标，MEF2 是序列特异性 DNA 结合蛋白。脊椎动物有 4 个 MEF2 家族成员，它们参与多种发育和信号转导过程，包括心脏发育。

HDAC 不仅参与特定基因的活动，而且也参与整体性的过程。酿酒酵母的 Rpd3 和 Hda1 也参与整体组蛋白去乙酰化（Vogelauer et al. 2000）。另外，HDAC 参与异染色质组装。酿酒酵母类型 3 HDAC 的 Sir2 在一个复合物中与 Sir4 相互作用，Sir4 通过序列特异性因子 Rap1 靶向到酿酒酵母的异染色质上（Grunstein 1997）。去乙酰化为 Sir3 的结合产生一个表面，然后 Sir3 结合 Sir2 和 Sir4。之后通过 Sir2 使邻近的核小体继续

去乙酰化，继而使去乙酰化扩展到一个区域，从而形成异染色质结构。

H3K9 甲基化和 HP1

臂间异染色质或着丝点周围异染色质的主要赖氨酸修饰之一是 H3K9 的三甲基化（Martin and Zhang 2005，2007；Shilatifard 2006；Kouzarides 2007）。正如上面所讨论的，Suv39Hl 和 Suv39H2 组蛋白甲基化转移酶催化这种修饰。Suv39 第一次在果蝇中被分离为位置效应花斑的抑制蛋白，称为 Su（Var）3-9。异染色质富集 H3K4me3，而常染色质富集 H3K9me2，但常染色质也包含一些 H3K9me3。高度相关的 G9a 和 GLP 蛋白的异源二聚体是主要的常染色质 H3K9 二甲基化酶。SetDB 或 ESET 是在异染色质和常染色质中发现的 H3K9 三甲基化酶，并且以自我调节的方式靶向到特定的锌指基因上，正如很多锌指抑制因子通过称为 KAP-1 的共抑制因子募集 SetDB 一样（见第 10 章）。

H3K9me2 和 H3K9me3 充当异染色质蛋白 1（HP1）家族蛋白的结合位点（图 1.15）（Lomberk et al. 2006）。称为 α、β 及 γ 的三个 HP1 家族成员存在于哺乳动物细胞中。虽然在缺少 H3K9 甲基化的酿酒酵母中没有发现 HP1，但 HP1 发现于裂殖酵母和很多其他真核细胞中，在裂殖酵母中称之为 Swi6。HP1 含有三个保守区：氨基末端染色域、可变连接子，以及羧基末端染色影结构域。染色域通过疏水口袋直接结合 H3K9me2 和 H3K9me3（Taverna et al. 2007）。染色影结构域在结构上与染色域相似，但缺乏对结合甲基化组蛋白必不可少的残基。反之，染色影结构域使 HP1 家族成员二聚化，这对染色质的 HP1 结合非常重要。

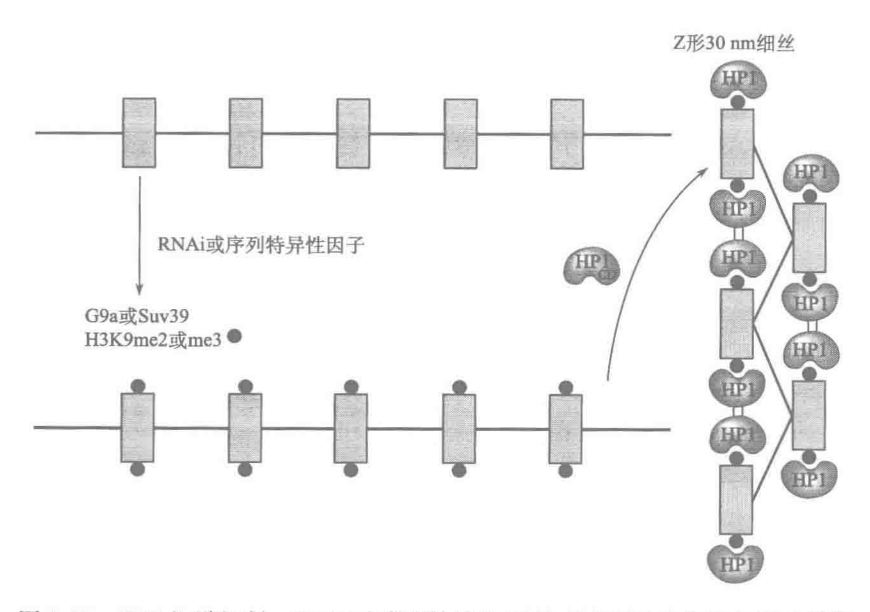


图 1.15　HP1 沉默机制。RNAi 或序列特异性 DNA 结合蛋白在异染色质中募集 Suv39 或在常染色质中募集 G9a。Suv39 三甲基化 H3K9，而 G9a 二甲基化 H3K9，这种修饰为 HP1 染色域（CD）提供一个结合位点。HP1 结合二甲基化和三甲基化的 H3K9。注意，每个核小体都有两个被修饰残基。HP1 染色影结构域与邻近核小体的染色影结构域二聚化，从而使压缩 Z 形双头螺旋形成。

二聚体化的染色影结构域也同参与异染色质组装的很多蛋白质，包括 SUV39 相互作用。事实上，认为 SUV39-HP1 相互作用为异染色质扩展提供了基本的机制，类似于酿酒酵母中 Sir2/4 相互作用（Martin and Zhang 2005）。HP1 也与共抑制因子，如 KAP1/TIF113 相互作用，KAP1/TIF113 通过 PXVXL 模体结合 HP1。值得注意的是，染色影也与 TAF4 相互作用；TAF4 是 TFIID 特异的 TAF。连接子区是 HP1 最进化分歧的片段，并与 DNA 和 RNA 相互作用（这一点我们在下面讨论）。

正如光漂白荧光恢复技术（FRAP）所测定的，HP1 在体外在微摩尔浓度范围内结合 H3K9me2 和 H3K9me3，并且在体内与异染色质的结合高度活跃（Schmiedeberg et al. 2004）。HP1α 和 HP1β 最初定位在臂间异染色质，而 HP1γ 定位在异染色质和常染色质上。这就是说，很多基因被认为在常染色质中受 HP1α 和 HP1β 的特定调控（Martin and Zhang 2005）。正如上述所描述的，HP1γ 是在活跃转录基因中发现的。HP1 在体外较喜欢结合包含 H2AZ 的核小体阵列，并有助于这些阵列形成 Z 形结构特征的压缩染色质。HP1 二聚化对压缩结构的形成是必需的。

还不完全清楚哺乳动物细胞中 H3K9 甲基化是怎样被靶向的。G9a 或 GLP 的敲除导致 H3K9me2 在常染色质沉默区中的急剧减少（Rice et al. 2003）。因此，已经提出在抑制期间 G9a 和 GLP 被序列特异性结合蛋白募集到常染色质基因上，并且已经知道了少数这种募集的实例。相比之下，很多锌指 DNA 结合蛋白含有同 KAP-1 相互作用的 KRAB 结构域。KAP-1 依次结合 SetDB1（见第 10 章）。该事件导致 H3K9 的三甲基化和 HP1 的募集。

关于 H3K9 三甲基化在裂殖酵母和果蝇中怎样靶向到异染色质，已有了一些认识（Grewal and Elgin 2007；Grewal and Jia 2007）。在裂殖酵母中的研究揭示 RNA 干扰（RNAi）通路对 HP1 结合和异染色质形成是必需的，并且相关的证据提示这种现象对于果蝇也可能是真实的。自相矛盾的是，在这些生物体中的研究暗示 DNA（通常是重复 DNA）的活跃转录对沉默来说是必需的。RNA 诱导转录沉默复合物（RITS）和 RNA 指导 RNA 聚合酶（RDRP）这两种复合物对裂殖酵母中的 H3K9 三甲基化是关键的。RITS 通过其 argonaute 亚基 ago1 同小干扰 RNA（siRNA）结合。这些 siRNA 可能将复合物指引到来自着丝粒异染色质重复的小转录本上，并且将这些转录本加工成 siRNA，而这些转录本依靠 RDRP 的复制似乎是非常重要的。RITS 怎样募集 Clr4 还不清楚，Clr4 与 Suv39 同源。然而，在裂殖酵母中，RITS、Clr4 及 Swi6（HP1）共定位于大多数异染色质上。哺乳动物细胞缺乏 RDRP，但含有来自异染色质重复的小转录本。除 RNAi 外，某些结合重复元件、交配型区域，以及着丝点的序列特异性蛋白与 HP1 的募集有关。

H3K27 甲基化和多梳蛋白

通过多梳蛋白 EZH2 的 H3K27 三甲基化仍然是沉默的另一种机制，估计对 4% 或更少的人类基因产生作用（Cheng et al. 2005；Daniel et al. 2005；Martin and Zhang 2005；Shilatifard 2006；Martin and Zhang 2007）。EZH2 首先在果蝇中被表征为一种发育的胚期中 *Hox* 基因的稳定沉默所必需的蛋白质。EZH2 是组蛋白甲基化酶的 SET 家族的创始成员之一。EZH2 同 EED、SUZ12、RbAP46 及 RbAP48 一起是多梳蛋白抑制

复合物 2（PRC2）的一个亚基。果蝇和哺乳动物中另一个 PRC 复合物称为 PRC1，它包括 Ring1B、Ring1A、HPC 1～3（人多梳蛋白或人 Pc）、HPH 1～3、Bmi1、Mell8 及 ScmH1～2。

HPC 含有一个染色域，同 HP1 的染色域非常相似。然而，结构学研究表明，微小的差异使它识别 H3K27me3 而不是 H3K9me3。在最简单的抑制模式中，PRC2 通过 DNA 结合蛋白靶向到染色质，并在此使 PRC2 三甲基化 H3K27。这种修饰转而招募 PRC1，尽管有可能存在其他因子促成结合。已知 PRC1 在体内同核小体形成稳定的抑制结构，这种结构阻止 SWI/SNF 重塑复合物的访问。PRC1 亚基 Ring1A 与 Ring1B 和 Bmi1 一起具有靶向于 H2AK119 的组蛋白泛素化活性。

在果蝇中，多梳蛋白抑制元件（PRE）募集 PRC 复合物到基因（Schuettengruber et al. 2007）。尽管该领域发展迅速，但在这一阶段，哺乳动物中没有已知的 PRE 存在。如果它们存在，将会被很快鉴定出来。一直以来，假定很多序列特异性结合蛋白在果蝇中募集 PRC2。在果蝇中，序列特异性结合蛋白 Pho 结合并募集 PRC2 到控制 Ubx 基因的 PRE 上。然而，它也独立募集 PRG1。

和研究者可能的预期相反，果蝇中 PRE 经常离启动子数千碱基，其抑制机制并不清楚，但是它可能参与染色质环的形成。另外，认为 PRE 很少有核小体，从而在某些情况下 PRC2 修饰通过 HPC 亚基募集 PRC1 的想法似乎不太可能。已经表明，RNAi 机器与非编码 RNA（如 hotair 和 Xists）一起参与 PRC2 及数个其他序列特异性 DNA 结合蛋白（如 GAGA 因子和 Pho）的募集。

全基因组研究显示了在 H3K27 上被三甲基化并与 Pol II 的出现呈负相关的大的基因组区域（Ringrose 2007；Ringrose and Paro 2007），即使某些失活基因含有暂停的 Pol II。最后，哺乳动物中的全基因研究显示，在沉默基因上，PRC1 和 PRC2 复合物之间有相当大的重叠，说明这两种复合物协同工作，支持一个模式，在该模式中 PRC2 修饰染色质以便 PRC1 结合。然而，它们也可能独立结合在某些位置。

CpG 上的 DNA 甲基化

在哺乳动物，而不是果蝇或酿酒酵母（至少到任何显著程度）中，在序列 CpG 的大沟中的胞嘧啶 C5 位置上，DNA 可被甲基化（Klose and Bird 2006；Martin and Zhang 2007；Miranda and Jones 2007）。CpG 甲基化在基因沉默中发挥作用。CpG 甲基化由称为 DNMT1、DNMT3A 及 DNMT3B 的三个主要甲基化酶催化，三种酶以 SAM 为辅因子，以便将甲基赠予胞嘧啶。DNMT3L 结合 DNMT3A 和 DNMT3B 并刺激它们的活性。有两种类型的 DNMT，即维持 DNMT 和从头 DNMT。据认为，DNMT1 是主要的维持甲基化酶，并在 DNA 复制叉处将半甲基化的 DNA 转变成全甲基化 DNA。DNMT3A 和 DNMT3B 是从头甲基化转移酶，并被募集到基因致使基因沉默。带有 CpG 岛的启动子正常情况下不被甲基化，但在某些疾病（如癌症）状态下它们被甲基化（Miranda and Jones 2007）。异常甲基化是癌症发展期间肿瘤抑制基因变得沉默的一种方式。另外，CpG 富集的启动子在失活 X 染色体上被甲基化。因为 DNA 甲基化可遗传，故认为 DNA 甲基化是一种比抑制性组蛋白甲基化更有效的沉默机制。事实上，抑制性组蛋白甲基化似乎同 DNA 甲基化在很多情况下有联系，说明一种功能性的相互关

系，其中与组蛋白修饰结合的蛋白质（如 HP1 和 PRC1）可能募集 DNMT，反之亦然。据认为某些因子，如 Myc 和早幼粒细胞白血病维甲酸受体-α（PML-RARα）融合，直接募集 DNMT3。同时还认为 RNAi 机器可能有助于募集 DNMT。

数种模式描述 DNA 甲基化怎样导致抑制（Klose and Bird 2006）。在一个模式中，甲基化抑制转录因子结合到因子结合位点的大沟。一小部分因子，包括 Igf2/H19 的印迹基因座中差异甲基化区域（DMR）上的绝缘子蛋白 CTCF，已在体内和体外表明了这一点。在第二个模式中，CpG 甲基化被甲基结合蛋白识别，甲基结合蛋白募集 HDAC 复合物（Fatemi and Wade 2006；Clouaire and Stancheva 2008）。

当前有 5 种蛋白质包含保守的 75 个氨基酸的甲基结合结构域（MBD）（Fatemi and Wade 2006；Clouaire and Stancheva 2008），这些蛋白质被称为 MeCP2、MBD1、MBD2、MBD3 及 MBD4。在哺乳动物中，MBD3 包含一个阻止结合甲基化 CpG 的突变，但是这个突变在其他脊椎动物中不存在（如非洲爪蟾属）。除包含 MBD 的蛋白质外，蛋白质的第二家族，即 Kaiso 家族，似乎通过不同的模体识别甲基 CpG。这个家族有三个成员，三个成员都有三个 C2H2 锌指，其中的两个锌指是结合甲基化 DNA 所必需的。然而，结合 Kaiso 的非甲基化序列也已被发现，表明这个家族有双重作用。

哺乳动物中大部分的 CpG 被甲基化，这引起一个问题，即 MBD 是否显示一些特异性。敲除小鼠的表型存在细微但显著的不同。MBD2 和 MeCP2 体内结合的 DNA 的 ChIP-Seq 分析揭示了不相重叠的特异性。虽然这可能产生于 MBD2 和 MeCP2 的独特的蛋白质-蛋白质相互作用，但是这两个蛋白质都包含可能识别 DNA 和赋予选择性的其他结构域。例如，MeCP2 有两个 AT-钩（AT-hook）结构域，该结构域结合 AT 富集区，并发现于染色质重塑蛋白和 HMGA 家族蛋白中。的确，MeCP2 好像结合邻近 AT 富集 DNA 的甲基化 CpG。已经知道了几个 MBD 的溶液结构（Clouaire and Stancheva 2008），并且最近已经解析出了带有甲基化 DNA 的 MeCP2 的晶体结构（Ho et al. 2008）。

关键的问题包括这些甲基结合蛋白怎样抑制转录和（或）它们是否参与其他过程。MeCP2 可能是研究得最广泛的甲基结合蛋白家族成员。关于 MeCP2 是否是一个更大的复合物的一部分，不同的研究导致矛盾的结果。然而，无论哪种方式，已经很明确 MeCP2 与 Sin3A/HDAC2 复合物及 Brm1 型哺乳动物 SWI/SNF 的募集有关。正如上面所描述的，MBD2 是 Mi-2/NuRD 复合物的一个亚基。另外，已发现 SETDB1 和 SUV39 同 MDB1 有关，以及 Kaiso 与 NcoR 共纯化。因此，甲基化 CpG 结合蛋白同共抑制复合物肯定有关系。最后，MeCP2 在体外染色质压缩处高度有效。在体外，MeCP2 结合核小体的出入点，并且发生多聚体化以形成压缩结构，该结构不同于由 HP1 促成的 Z 形结构。自相矛盾的是，MeCP2 只是一个抑制因子的思想最近受到挑战，因为在小鼠下丘脑中活跃转录的基因上发现了 MeCP2，而在 MeCP2 敲除小鼠中这些基因的活性下降（Cohen et al. 2008）。

结语

在过去的 10 年间，技术的进步已导致关于真核基因调控信息的激增。从上面梗概

中显而易见，我们对转录机器和修饰已经所知甚多；然而，除一些基本事件（如 Pol II 起始）以外，精确的机制还不清楚。关于将染色质同基因控制相联系的事件，已出现了数量惊人的描述性信息。然而，重要的是要认识到这些观察资料代表一种构想出假说的途径。后面各章节的精神是为这一领域的研究者提供策略和概念，以便他们利用科学的方法开始严格地测试他们的想法。

参考文献

Albright S. R. and Tjian R. 2000. TAFs revisited: More data reveal new twists and confirm old ideas. *Gene* 242: 1-13.

Allis C. D., Berger S. L., Cote J., Dent S., Jenuwien T., Kouzarides T., Pillus L., Reinberg D., Shi Y., Shiekhattar R., et al. 2007. New nomenclature for chromatin-modifying enzymes. *Cell* 131: 633-636.

Anand R. and Marmorstein R. 2007. Structure and mechanism of lysine-specific demethylase enzymes. *J. Biol. Chem.* 282: 35425- 35429.

Andel F., 3rd, Ladurner A. G., Inouye C., Tjian R., and Nogales E. 1999. Three-dimensional structure of the human TFIID-IIA-IIB complex. *Science* 286: 2 153-2156.

Armstrong J. A. 2007. Negotiating the nucleosome: Factors that allow RNA polymerase II to elongate through chromatin. *Biochem. Cell Biol.* 85: 426-434.

Bao Y. and Shen X. 2007a. INO80 subfamily of chromatin remodeling complexes. *Mutat. Res.* 618: 18-29.

Bao Y. and Shen X. 2007b. Snapshot: Chromatin remodeling complexes. *Cell* 129: 632.

Barski A., Cuddapah S., Cui K., Roh T. Y., Schones D. E., Wang Z., Wei G., Chepelev I., and Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* 129: 823-837.

Bernstein E. and Hake S. B. 2006. The nucleosome: A little variation goes a long way. *Biochem. Cell Biol.* 84: 505-517.

Biddick R. and Young E. T. 2005. Yeast mediator and its role in transcriptional regulation. *C. R. Biol.* 328: 773-782.

Bortvin A. and Winston F. 1996. Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science* 272: 1473-1476.

Bourbon H. M. 2008. Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. *Nucleic Acids Res.* 36: 3993-4008.

Branco M. R. and Pombo A. 2006. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol.* 4: e138.

Cal Y., Jin J., Yao T., Gottschalk A. J., Swanson S. K., Wu S., Shi Y., Washburn M. P., Florens L., Conaway R. C., and Conaway J. W. 2007. YY1 functions with INO80 to activate transcription. *Nat. Struct. Mol. Biol.* 14: 872-874.

Cairns B. R. 2005. Chromatin remodeling complexes: Strength in diversity, precision through specialization. *Curr. Opin. Genet. Dev.* 15: 185-190.

Carey M. 1998. The enhanceosome and transcriptional synergy. *Cell* 92: 5-8.

Carter D. R., Eskiw C., and Cook P. R. 2008. Transcription factories. *Biochem. Soc. Trans.* 36: 585-589.

Chadick J. Z. and Asturias F. J. 2005. Structure of eukaryotic Mediator complexes. *Trends Biochens. Sci.* 30: 264-271.

Chakaiova L., Carter D., and Fraser P. 2004. RNA fluorescence in situ hybridization tagging and recovery of associated proteins to analyze in vivo chronsatin interactions. *Methods Enzyrnol* 375: 479-493.

Chakravarthy S., Park Y. J., Chodaparambil J., Edayathumangalam R. S., and Luger K. 2005. Structure and dynamic properties of nucleosome core particles. *FEBS Lett.* 579: 895-898.

Chédin S., Laferté A. Hoang T., Lafontaine D. L., Riva M., and Caries C. 2007. Is ribosome synthesis controlled by PolI transcription? *Cell Cycle* 6: 11-15.

Chen B. S. and Hampsey M. 2002. Transcription activation: Unveiling the essential nature of TFIID. *Curr. Biol.* 12: R620-R622.

Chen H. T. and Hahn S. 2003. Binding of TFIIB to RNA polymerase II: Mapping the binding site for the TFIIB zinc ribbon domain within the preinitiation complex. *Mol. Cell* 12: 437-447.

Chen H. T. and Hahn S. 2004. Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: A model for the structure of the PIC. *Cell* 119: 169-180.

Chen M. and Shen X. 2007. Nuclear actin and actin-related proteins in chromatin dynamics. *Curr. Opin. Cell Blot.* 19: 326-330.

Cheng X., Collins R. E., and Zhang X. 2005. Structural and sequence motifs of protein (histone) methylation enzymes. *Annu. Rev. Biophys. Biomol. Struct.* 34: 267-294.

Chi T., Liebernsan P., Ellwood K., and Carey M. 1995. A general mechanism for transcriptional synergy by eukaryotic activators. *Nature* 377: 254-257.

Clouaire T. and Stancheva I. 2008. Methyl-CpG binding proteins: Specialized transcriptional repressors or structural components of chromatin? *Cell. Mat. Life Sci.* 65: 1509-1522.

Cohen S., Zhou Z., and Greenberg M. E. 2008. Medicine. Activating a repressor. *Science* 320: 1172-1173.

Collas P. and Dahl J. A. 2008. Chop it, ChIP it, check it: The current status of chromatin immunoprecipitation. *Front. Biosci.* 13: 929-943.

Conaway R. C., Sato S., Tomomori-Sato C., Yao T., and Conaway J. W. 2005. The mammalian Mediator complex and its role in transcriptional regulation. *Trends Biochem. Sci.* 30: 250-255.

Cosma M. P. 2002. Ordered recruitment: Gene-specific mechanism of transcription activation. *Mol. Cell* 10: 227-236.

Cramer P., Armache K. J., Baumli S., Benkert S., Brueckner F., Buchen C., Damsma G. E., Dengl S., Geiger S. R., Jasiak A. J., et al. 2008. Structure of eukaryotic RNA polymerases. *Annu. Rev. Biophys.* 37: 337-352.

Crawford G. E., Davis S., Scacheri P. C., Renaud G., Halawi M. J., Erdos M. R., Green R., Meltzer P. S., Wolfsberg T. G., and Collins F. S. 2006a. DNase-chip: A high-resolution method to identify DNase I hypersensitive sites using tiled microarrays. *Nat. Methods* 3: 503-509.

Crawford G. E., Holt I. E., Whittle j., Webb B. D., Tai D., Davis S., Margulies E. H., Chen Y., Bernat J. A., Ginsburg D., et al. 2006b. Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). *Genonse Res.* 16: 123-131.

Cremer T., Cremer M., Dietzel S., Muller S., Solovei I., and Fakan S. 2006. Chromosome territories-A functional nuclear landscape. *Curr. Opin. Cell Biol.* 18: 307-316.

Daniel J. A., Pray-Grant M. G., and Grant P. A. 2005. Effector proteins for methylated histones: An expanding family. *Cell Cycle* 4: 919-926.

Davidson I., Kobi D., Fadloun A., and Mengus G. 2005. New insights into TAFs as regulators of cell cycle and signaling pathways. *Cell Cycle* 4: 1486-1490.

Dean A. 2006. On a chromosome far, far away: LCRs and gene expression. *Trends Genet.* 22: 38-45.

Dechassa M. L., Zhang B., Horowitz-Scherer R., Persinger J., Woodcock C. L., Peterson C. L., and Bartholomew B. 2008. Architecture of the SWI/SNF-nucleosome complex. *Mol. Cell Biol.* 28: 6010-6012.

de Laat W. and Grosveld F. 2003. Spatial organization of gene expression: The active chromatin hub. *Chromosome Res.* 11: 447-459.

Dieci G., Fiorino G., Castelnuovo M., Teichmann M., and Pagano A. 2007. The expanding RNA polymerase III transcriptome. *Trends Genet.* 23: 614-622.

Ding N., Zhou H., Esteve P. O., Chin H. G., Kim S., Xu X., Joseph S. M., Friez M. J., Schwartz C. E., Pradhan S., and Boyer T. G. 2008. Mediator links epigenetic silencing of neuronal gene expression with X-linked mental retardation. *Mol. Cell* 31: 347-359.

Dorman E. R., Bushey A. M., and Corces V. G. 2007. The role of insulator elements in large-scale chromatin structure in interphase. *Semin. Cell Dev. Biol.* 18: 682-690.

Dyda F., Klein D. C., and Hickman A. B. 2000. GCN5-related N-acetyltransferases: A structural overview. *Annu. Rev. Biophys. Biomol. Struct.* 29: 81-103.

Eirín-López J. and Ausió J. 2007. H2A. Z-mediated genome-wide chromatin specialization. *Curr. Genomics* 8: 59-66.

Eissenberg J. C. and Shilatifard A. 2006. Leaving a mark: The many footprints of the elongating RNA polymerase II. *Curr. Opin. Genet. Dev.* 16: 184-190.

The ENCODE Project Consortium. 2004. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* 306: 636-640.

Fatemi M. and Wade P. A. 2006. MBD family proteins: Reading the epigenetic code. *J. Cell Sci.* 119: 3033-3037.

Fazzio T. G., Kooperberg C., Goldmark J. P., Neal C., Basom R., Delrow J., and Tsukiyama T. 2001. Widespread collaboration of Isw2 and Sin3-Rpd3 chromatin remodeling complexes in transcriptional repression. *Mol. Cell. Biol.* 21: 6450-6460.

Ferrari R., Pellegrini M., Horwitz G. A., Xie W., Berk A. J., and Kurdistani S. K. 2008. Epigenetic reprogramming by adenovirus ela. *Science* 321: 1086-1088.

Ferreira H., Flaus A., and Owen-Hughes T. 2007. Histone modifications influence the action of Snf2 family remodelling enzymes by different mechanisms. *J. Mol. Biol.* 374: 563-579.

Fu Y., Sinha M., Peterson C. L., and Weng Z. 2008. The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet.* 4: e1000138.

Galande S., Purbey P. K., Notani D., and Kumar P. P. 2007. The third dimension of gene regulation: Organization of dynamic chromatin loopscape by SATB1. *Curr. Opin. Genet. Dev.* 17: 408-414.

Grewal S. I. and Elgin S. C. 2007. Transcription and RNA interference in the formation of heterochromatin. *Nature* 447: 399-406.

Grewal S. I. and Jia S. 2007. Heterochromatin revisited. *Nat. Rev. Genet.* 8: 35-46.

Grob P., Cruse M. J., Inouye C., Peris M., Penczek P. A., Tjian R., and Nogales E. 2006. Cryo-electron microscopy studies of human TFIID: Conformational breathing in the integration of gene regulatory cues. *Structure* 14: 511-520.

Grunstein M. 1997. Molecular model for telomeric heterochromatin in yeast. *Curr. Opin. Cell Bull.* 9: 383-387.

Haeusler R. A. and Engelke D. R. 2006. Spatial organization of transcription by RNA polymerase III. *Nucleic Acids Res.* 34: 4826-4836.

Hahn S. 1998. The role of TAFs in RNA polymerase II transcription. *Cell* 95: 579-582.

Hahn S. 2004. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* 11: 394-403.

Hake S. B. and Allis C. D. 2006. Histone H3 variants and their potential role in indexing mammalian genomes: The "H3 barcode hypothesis." *Proc. Natl. Acad. Sci.* 103: 6428-6435.

Hall J. A. and Georgel P. T. 2007. CHD proteins: A diverse family with strong ties. *Biochem. Cell Biol.* 85: 463-476.

Hatzis P. and Talianidis I. 2002. Dynamics of enhancer-promoter communication during differentiation-induced gene activation. *Mol. Cell* 10: 1467-1477.

Heintzman N. D., Stuart R. K., Hon G., Fu Y., Ching C. W., Hawkins R. D., Barrera L. O., Van Calcar S., Qu C., Ching K. A., et al. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* 39: 311-318.

Ho K. L., McNae I. W., Schmiedeberg L., Klose R. J., Bird A. P., and Walkinshaw M. D. 2008. MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol. Cell* 29: 525-531.

Hoch D. A., Stratton J. J., and Gloss L. M. 2007. Protein-protein Forster resonance energy transfer analysis of nucleosome core particles containing H2A and H2A. Z. *J. Mol. Biol.* 371: 971-988.

Hock R., Furusawa T., Ueda T., and Bustin M. 2007. HMG chromosomal proteins in development and disease. *Trends Cell Biol.* 17: 72-79.

Hoffmann A., Chiang C. M., Oelgeschlager T., Xie X., Burley S. K., Nakatani Y., and Roeder R. G. 1996. A histone octamer-like structure within TFIID. *Nature* 380: 356-359.

Holstege F. C., Tantin D., Carey M., van der Vliet P. C., and Timmers H. T. 1995. The requirement for the basal transcription factor TIE is determined by the helical stability of promoter DNA. *EMBO J.* 14: 810-819.

Horn P. J. and Peterson C. L. 2002. Molecular biology. Chromatin higher order folding-Wrapping up transcription. *Science* 297: 1824-1827.

Jacobson R. H., Ladurner A. G., King D. S., and Tjian R. 2000. Structure and function of a human TAFII250 double bromodomain module. *Science* 288: 1422-1425.

Jin J., Cal Y., Li B., Conaway R. C., Workman J. L, Conaway J. W., and Kusch T. 2005. In and out: Histone variant exchange in chromatin. *Trends Biochem. Sci.* 30: 680-687.

Johnson K. M., Wang J., Smallwood A., Arayata C., and Carey M. 2002. TFIID and human mediator coactivator complexes assemble cooperatively on promoter DNA. *Genes Dev.* 16: 1852-1863.

Kadauke S. and Blobel G. A. 2008. Chromatin loops in gene regulation. *Biochim. Biophys. Acta* (in press).

Kaplan C. D., Laprade L., and Winston F. 2003. Transcription elongation factors repress transcription initiation from cryptic Sites. *Science* 301: 1096-1099.

Kim T. H. and Ren B. 2006. Genome-wide analysis of protein-DNA interactions. *Annu. Rev. Genomics Hum. Genet.* 7: 81-102.

Kim T. K., Ebright R. H., and Reinberg D. 2000. Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* 288: 1418-1422.

Kim T. H., Abdullaev Z. K., Smith A. D., Ching K. A., Loukinov D. I., Green R. D., Zhang M. Q., Lobanenkov V. V., and Ren B. 2007. Analysis of the vertebrate insulator protein CTCF-binding Sites in the human genome. *Cell* 128: 1231-1245.

Kim T. H., Barrera L. O., Zheng M., Qu C., Singer M. A., Richmond T. A., Wu Y., Green R. D., and Ren B. 2005. A high-resolution map of active promoters in the human genome. *Nature* 436: 876-880.

Klose R. J. and Bird A. P. 2006. Genomic DNA methylation: the mark and its mediators. *Trends Biochem. Sci.* 31: 89-97.

Konev A. Y., Tribus M., Park S. Y., Podhraski V., Lim C. Y., EmelyanovA. V., Vershiova E., Pirrotta V., Kadonaga J. T., Lusser A., and Fyodorov D. V. 2007. CHD1 motor protein is required for deposition of histone variant H3. 3 into chromatin in vivo. *Science* 317: 1087-1090.

Korber P. and Hörz W. 2004. SWRred not shaken: Mixing the histones. *Cell* 117: 5-7.

Kornberg R. D. 2005. Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.* 30: 235-239.

Kornberg R. D. 2007. The molecular basis of eukaryotic transcription. *Proc. Natl. Acad. Sci.* 104: 12955-12961.

Kouzarides T. 2007. Chromatin modifications and their function. *Cell* 128: 693-705.

Kurdistani S. K., Tavazoie S., and Grunstein M. 2004. Mapping global histone acetylation patterns to gene expression. *Cell* 117: 721-733.

Kusch T. and Workman J. L. 2007. Histone variants and complexes involved in their exchange. *Subcell. Biochem.* 41: 91-109.

Lakowski B., Roelens I., and Jacob S. 2006. CoREST-like complexes regulate chromatin modification and neuronal gene expression. *J. Mol. Neurosci.* 29: 227-239.

Lan F., Nottke A. C., and Shi Y. 2008. Mechanisms involved in the regulation of histone lysine demethylases. *Curr. Opin. Cell Biol.* 20: 316-325.

Lee K. K. and Workman J. L. 2007. Histone acetyltransferase complexes: One size doesn't fit all. *Nat. Rev. Mol. Cell Biol.* 8: 284-295.

Lefstin J. A. and Yamamoto K. R. 1998. Allosteric effects of DNA on transcriptional regulators. *Nature* 392: 885-888.

LeRoy G., Rickards B., and Flint S. J. 2008. The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. *Mol. Cell* 30: 51-60.

Leschziner A. E., Saha A., Wittmeyer J., Zhang Y., Bustamante C., Cairns B. R., and Nogales E. 2007. Conformational flexibility in the chromatin remodeler RSC observed by electron microscopy and the orthogonal tilt reconstruction method. *Proc. Natl. Acad. Sci.* 104: 4913-4918.

Leurent C., Sanders S., Ruhlmann C., Mallouh V., Weil P. A., Kirschner D. B., Tora L., and Schultz P. 2002. Mapping histone fold TAFs within yeast TFIID. *EMBO J.* 21: 3424-3433.

Li B., Carey M., and Workman J. L. 2007a. The role of chromatin during transcription. *Cell* 128: 707-719.

Li A. G., Piluso L. G., Cai X., Gadd B. J., Ladurner A. G., and Liu X. 2007b. An acetylation switch in p53 mediates holo-TFIID recruitment. *Mol. Cell* 28: 408-421.

Li B., Gogol M., Carey M., Lee D., Seidel C., and Workman J. L. 2007c. Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. *Science* 316: 1050-1054.

Li B., Gogol M., Carey M., Pattenden S. G., Seidel C., and Workman J. L. 2007d. Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. *Genes Dev.* 21: 1422-1430.

Linnemann A. K., Platts A. E., Doggett N., Gluch A., Bode J., and Krawetz S. A. 2007. Genomewide identification of nuclear matrix attachment regions: An analysis of methods. Biochem. *Soc. Trans.* 35: 612-617.

Liu Y., Kung C., Fishburn J., Ansari A. Z., Shokai K. M., and Hahn S. 2004. Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. *Mol. Cell Biol.* 24: 1721-1735.

Lomberk G., Wallrath L., and Urrutia R. 2006. The heterochromatin protein 1 family. *Genome Biol.* 7: 228.

Lorch Y., Maier-Davis B., and Kornberg R. D. 2006. Chromatin remodeling by nucleosome disassembly in vitro. *Proc. Natl. Acad. Sci.* 103: 3090-3093.

Loyola A. and Almouzni G. 2007. Marking histone H3 variants: How, when and why? *Trends Biochem. Sci.* 32: 425-433.

Luger K. and Hansen J. C. 2005. Nucleosome and chromatin fiber dynamics. *Curr. Opin. Struct. Biol.* 15: 188-196.

Maeda R. K. and Karch F. 2007. Making connections: Boundaries and insulators in *Drosophila*. *Curr. Opin. Genet. Dev.* 17: 394-399.

Mahajan M. C., Karmakar S., and Weissman S. M. 2007. Control of β globin genes. *J. Cell. Biochem.* 102: 801-810.

Malik S. and Roeder R. G. 2005. Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem. Sci.* 30: 256-263.

Margaritis T. and Holstege F. C. 2008. Poised RNA polymerase II gives pause for thought. *Cell* 133: 581-584.

Martin C. and Zhang Y. 2005. The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* 6: 838-849.

Martin C. and Zhang Y. 2007. Mechanisms of epigenetic inheritance. *Curr. Opin. Cell Biol.* 19: 266-272.

Maston G. A., Evans S. K., and Green M. R. 2006. Transcriptional regulatory elements in the human genome. *Annu. Rev. Genomtcs Hum. Genet.* 7: 29-59.

Matangkasombut O., Buratowski R. M., Swilling N. W., and Buratowski S. 2000. Bromodomain factor 1 corresponds to a missing piece of yeast TFIID. *Genes Dev.* 14: 951-962.

McBryant S. J. , Adams V. H. , and Hansen J. C. 2006. Chromatin architectural proteins. *Chromosome Res*. 14: 39-51.

Mellor J. and Morillon A. 2004. ISWI complexes in *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* 1677: 100-112.

Merika M. and Thanos D. 2001. Enhanceosomes. *Curr. Opin. Genet. Dev.* 11: 205-208.

Mihardja S. , Spakowitz A. J. , Zhang Y. , and Bustamante C. 2006. Effect of force on mononucleosomal dynamics. *Proc. Natl. Acad. Sci.* 103: 15871-15876.

Miller G. and Hahn S. 2006. A DNA-tethered cleavage probe reveals the path for promoter DNA in the yeast preinitiation complex. *Nat. Struct. Mol. Biol.* 13: 603-610.

Miranda T. B. and Jones P. A. 2007. DNA methylation: The nuts and bolts of repression. *J. Cell Physiol*. 213: 384-390.

Mitchell J. A. and Fraser P. 2008. Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev*. 22: 20-25.

Mohrmann L. and Verrijzer C. P. 2005. Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. *Biochim. Biophys. Acta* 1681: 59-73.

Mujtaba S. , Zeng L. , and Zhou M. M. 2007. Structure and acetyllysine recognition of the bromodomain. *Oncogene* 26: 5521-5527.

Oelgeschlager T. , Chiang C. M. , and Roeder R. G. 1996. Topology and reorganization of a human TFIID-promoter complex. *Nature* 382: 735-738.

Paoletti A. C. , Parmely T. J. , Tomomori-Satd C. , Sato S. , Zhu D. , Conaway R C. , Conaway J. W. , Florens L. , and Washburn M. P. 2006. Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc. Natl. Acad. Sci.* 103: 18928-18933.

Parelbo V. , Hadjur S. , Spivakov M. , Leleu M. , Sauer S. , Gregson H. C. , Jarmuz A. , Canzonetta C. , Webster Z. , Nesterova T. , et al. 2008. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 132: 422-433.

Park Y. J. and Luger K. 2006. Structure and function of nucleosome assembly proteins. *Biochem. Cell Biol*. 84: 549-558.

Pavri R. , Zhu B. , Li G. , Trojer P. , Mandal S. , Shilatifard A. , and Reinberg D. 2006. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* 125: 703-717.

Peterlin B. M. and Price D. H. 2006. Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell* 23: 297-305.

Peterson C. L. and Laniel M. A. 2004. Histones and histone modifications. *Curr. Biol.* 14: R546-R551.

Peterson C. L. and Workman J. L. 2000. Promoter targeting and chromatin remodeling by the SWI/SNF complex. *Curt. Opin. Genet. Dev.* 10: 187-192.

Phatnani H. P. and Greenleaf A. L. 2006. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev*. 20: 2922-2936.

Ptashne M. and Gann A. A. 1990. Activators and targets. *Nature* 346: 329-331.

Qiu H. , Hu C. , Wong C. M. , and Hinnebusch A. G. 2006. The Spt4p subunit of yeast DSIF stimulates association of the Paf1 complex with elongating RNA polymerase II. *Mol. Cell. Biol.* 26: 3135-3148.

Reina J. H. and Hernandez N. 2007. On a roll for new TRF targets. *Genes Dev*. 21: 2855-2860.

Reinberg D. and Sims R. J. , 3rd. 2006. de FACTo nucleosome dynamics. *J. Biol. Chem.* 281: 23297-23301.

Reinke H. and Horz W. 2003. Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. *Mol. Cell* 11: 1599-1607.

Rhoades A. R. , Ruone S. , and Formosa T. 2004. Structural features of nucleosomes reorganized by yeast FACT and its HMG box component, Nhp6. *Mol. Cell. Biol.* 24: 3907-3917.

Rice J. C., Briggs S. D., Ueberheide B., Barber C. M., Shabanowitz J., Hunt D. F., Shinkai Y., and Allis C. D. 2003. Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol. Cell* 12: 1591-1598.

Richmond T. J. 2006. Genomics: Predictable packaging. *Nature* 442: 750-752.

Ringrose L. 2007. Polycomb comes of age: Genome-wide profiling of target sites. *Curr. Opin. Cell Blol.* 19: 290-297.

Ringrose L. and Paro R. 2007. Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development* 134: 223-232.

Roeder R. G. 2005. Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett.* 579: 909-915.

Rob T. Y., Wei G., Farrell C. M., and Zhao K. 2007. Genome-wide prediction of conserved and nonconserved enhancers by histone acetylation patterns. *Genome Res.* 17: 74-81.

Russell J. and Zomerdijk J. C. 2006. The RNA polymerase I transcription machinery. *Biochem. Soc. Symp.* 73: 203-216.

Saha A., Wittmeyer J., and Cairns B. R. 2006a. Chromatin remodelling: The industrial revolution of DNA around histones. *Nat. Rev. Mol. Cell Biol.* 7: 437-447.

Saha A., Wittmeyer J., and Cairns B. R. 2006b. Mechanisms for nudeosome movement by ATP-dependent chromatin remodeling complexes. *Results Probl. Cell Differ.* 41: 127-148.

Saunders A., Core L. J., and Lis J. T. 2006. Breaking barriers to transcription elongation. *Nat. Rev. Mol. Cell Biol.* 7: 557-567.

Scarpulla R. C. 2008. Transcriptional paradigms in mammalian mitochondrial biogenesis and function. *PhysioL Rev.* 88: 611-638.

Schalch T., Duda S., Sargent D. F., and Richmond T. J. 2005. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* 436: 138-141.

Schmiedeberg L., Weisshart K., Diekmann S., Meyer Zu Hoerste G., and Hemnierich P. 2004. High- and low-mobility populations of HP1 in heterochromatin of mammalian cells. *Mol. Biol Cell* 15: 2819-2833.

Schuettengruber B., Chourrout D., Vervoort M., Leblanc B., and Cavalli G. 2007. Genome regulation by polycomb and trithorax proteins. *Cell* 128: 735-745.

Segal E., Fondufe-Mittendorf Y., Chen L., Thastrom A., Field Y., Moore, I. K., Wang J. P., and Widom J. 2006. A genomic code for nucleosome positioning, *Nature* 442: 772-778.

Sexton T., Umlauf D., Kurukuti S., and Fraser P. 2007. The role of transcription factories in large-scale structure and dynamics of interphase chromatin. *Semin. Cell Dev. Biol.* 18: 691-697.

Shao H., Revach M., Moshonov S., Tzuman Y., Gazit K., Albeck S., Unger T., and Dikstein R. 2005. Core promoter binding by his- tone-like TAF complexes. *Mol. Cell. Biol.* 25: 206-219.

Shi Y. and Whetstine J. R. 2007. Dynamic regulation of histone lysine methylation by demethylases. *Mol. Cell* 25: 1-14.

Shilatifard A. 2006. Chromatin modifications by methylation and ubiquitination: Implications in the regulation of gene expression. *Annu. Rev. Biochem.* 75: 243-269.

Shilatifard A. 2008. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr. Opin. Cell Biol.* 20: 341-348.

Shilatifard A., Conaway R. C., and Conaway J. W. 2003. The RNA polymerase II elongation complex. *Annu. Rev. Biochem.* 72: 693-715.

Shogren-Knaak M., Ishii H., Sun J. M., Pazin M. J., Davie J. R., and Peterson C. L. 2006. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311: 844-847.

Silverstein R. A. and Ekwall K. 2005. Sin3: A flexible regulator of global gene expression and genome stability.

Curr. Genet. 47: 1-17.

Simonis M., Kooren J., and de Laat W. 2007. An evaluation of 3C-based methods to capture DNA interactions. *Nat. Methods* 4: 895-901.

Smale S. T. and Kadonaga J. T. 2003. The RNA polymerase II core promoter. *Annu. Rev. Biochem.* 72: 449-479

Smith C. L. 2008. A shifting paradigm: Histone deacetylases and transcriptional activation. *Bioessays* 30: 15-24.

Smith C. L., Horowitz-Scherer R., Flanagan J. F., Woodcock C. L., and Peterson C. L. 2003. Structural analysis of the yeast SWI/SNF chromatin remodeling complex. *Nat. Struct. Biol.* 10: 141-145.

Soutoglou E. and Talianidis I. 2002. Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation. *Science* 295: 1901-1904.

Struhl K. 2005. Transcriptional activation: Mediator can act after preinitiation complex formation. *Mol. Cell* 17: 752-754.

Svejstrup J. Q. 2007a. Contending with transcriptional arrest during RNAPII transcript elongation. *Trends Biochem. Sci.* 32: 165-171.

Svejstrup J. Q. 2007b. Elongator complex: How many roles does it play? *Curr. Opin. Cell Biol.* 19: 331-336.

Taatjes D. J., Marr M. T., and Tjian R. 2004a. Regulatory diversity among metazoan coactivator complexes. *Nat. Rev. Mol. Cell Biol.* 5: 403-410.

Taatjes D. J., Schneider-Poetsch T., and Tjian R. 2004b. Distinct conformational states of nuclear receptor-bound CRSP-Med complexes. *Nat. Struct. Mol. Biol.* 11: 664-671.

Taverna S. D., Li H., Ruthenburg A. J., Allis C. D., and Patel D. J. 2007. How chromatin-binding modules interpret histone modifications: Lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* 14: 1025-1040.

Tora L. 2002. A unified nomenclature for TATA box binding protein (TBP) -associated factors (TAFs) involved in RNA polymerase II transcription. *Genes dev.* 16: 673-675.

Torres-Padilla M. E. and Tora L. 2007. TBP homologues in embryo transcription: Who does what? *EMBO Rep.* 8: 1016-1018.

Treisman R. 1996. Regulation of transcription by MAP kinase cascades. *Curr. Opin. Cell Biol.* 8: 205-215.

Tremethick D. J. 2007. Higher-order structures of chromatin: The elusive 30 nm fiber. Cell 128: 65 1-654.

Trojer P. and Reinberg D. 2007. Facultative heterochromatin: Is there a distinctive molecular signature? *Mol. Cell* 28: 1-13.

Valenzuela L. and Kamakaka R. T. 2006. Chromatin insulators. *Annu. Rev. Genet.* 40: 107-138.

Vaughn M. W. and Martienssen R. A. 2005. Finding the right template: RNA Pol IV, a plant-specific RNA polymerase. *Mol. Cell* 17: 754-756.

Vermeulen M., Mulder K. W., Denissov S., Pijnappel W. W., van Schaik F. M., Varier R. A., Baltissen M. P., Stunnenberg H. G., Mann M., and Timmers H. T. 2007. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* 131: 58-69.

Vogelauer M., Wu J., Suka N., and Grunstein M. 2000. Global histone acetylation and deacetylation in yeast. *Nature* 408: 495-498.

Wallace J. A. and Felsenfeld G. 2007. We gather together: Insulators and genome organization. *Curr. Opin. Genet. Dev.* 17: 400-407.

Wang A., Kurdistani S. K., and Grunstein M. 2002. Requirement of Hos2 histone deacetylase for gene activity in yeast. *Science* 298: 1412-1414.

Wang G., Balamotis M. A., Stevens J. L., Yamaguchi Y., Handa H., and Berk A. J. 2005. Mediator requirement for both recruitment and postrecruitment steps in transcription initiation. *Mol. Cell* 17: 683-694.

Wang Q., Carroll J. S., and Brown M. 2005. Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. *Mol. Cell* 19: 631-642.

Wang Z., Zang C., Rosenfeld TA., Schones D. E., Barski A., Cuddapah S., Cui K., Roh T. Y., Peng W.,

Zhang M. Q., and Zhao K. 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* 40: 897-903.

Weake V. M. and Workman J. L. 2008. Clearing the way for unpaused polymeraaes. *Cell* 134: 16-18.

Whitehouse I., Rando O. J., Delrow J., and Tsukiyama T. 2007. Chromatin remodelling at promoters suppresses antisense transcription. *Nature* 450: 1031-1035.

Wisniewski J. R., Zougman A., Kruger S., and Mann M. 2007. Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations, and phosphorylation as well as differences between cell culture and tissue. *Mol. Cell. Proteomics* 6: 72-87.

Woodcock C. L. 2006. Chromatin architecture. *Curr. Opin. Struct. Biol.* 16: 213-220.

Woodcock C. L., Skoultchi A. I., and Fan Y. 2006. Role of linker histone in chromatin structure and function: H1 stoichiometry and nudeosome repeat length. *Chromosome Res.* 14: 17-25.

Workman J. L. 2006. Nucleosome displacement in transcription. *Genes Dev.* 20: 2009-2017.

Wu C., Bassett A., and Travers A. 2007. A variable topology for the 30-nm chromatin fibre. *EMBO Rep.* 8: 1129-1134.

Xi H., Shulha H. P., Lin J. M., Vales T. R., Fu V., Bodine D. M., McKay R. D., Chenoweth J. G., Tesar P. J., Furey T. S., et al. 2007. Identification and characterization of cell type-specific and ubiquitous chromatin regulatory structures in the human genome. *PLoS Genet.* 3: e136.

Xu L., Glass C. K., and Rosenfeld M. G. 1999. Coactivator and corepressor complexes in nuclear receptor function. *Curr. Opin. Genet. Dev.* 9: 140-147.

Yang C., Bolotin E., Jiang T., Sladek F. M., and Martinez E. 2007. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene* 389: 52-65.

Yang X. J. and Seto E. 2008. The Rpd3/Hdal family of lysine deacetylases: From bacteria and yeast to mice and men. *Nat. Rev. Mol. Cell Biol.* 9: 206-218.

Zhang H., Roberts D. N., and Cairns B. R. 2005. Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* 123: 219-231.

Zhao H. and Dean A. 2005. Organizing the genome: Enhancers and insulators. *Biochem. Cell Biol.* 83: 516-524.

Zilberman D. and Henikoff S. 2007. Genome-wide analysis of DNA methylation patterns. *Development* 134: 3959-3965.

2 初始策略性问题

要点

- 通过各种简单的实验可以对新转录因子的靶基因、转录因子参与的生物网络，以及因子调控基因表达的机制获得初步认识。然而，完成真正有意义的分析可能需要数年的功夫。
- 通过对潜在调控序列进行基本的计算分析，再结合体内外蛋白质与DNA相互作用分析，可获得目的基因调控机制的初步知识。然而，为了全面地理解这种策略，要求有更广泛的分析。

引言和概述

大多数研究人员都怀揣下列两个目标之一进入转录领域：①了解新鉴定的基因是如何调控的；②了解新转录因子的功能。由于各种各样的原因，可能会对新基因的调控机制产生兴趣。在细胞生物的发育过程中，基因可能以一种有趣的时间或空间模式表达，表明基因的调控因子发挥着重要的发育调控作用。此外，作为对刺激的应答，基因可能被激活或抑制；或者基因的异常调控可能在特定疾病中发挥作用。因此，基因的调控因子有助于理解疾病的发病机制或为治疗干预建议靶点。同样，可能对一个新的转录因子产生兴趣，因为其表达模式或突变表型暗示其在发育、细胞活化或疾病病理中发挥重要作用。

带着这些目标之一，许多研究人员已选择将大量的资源转向基因调控研究，这是令人钦佩的目标，因为他们的研究结果有助于我们理解控制重要细胞过程的调控线路。然而，在开始这种分析之前，必须慎重评估实现特定目标所需要的精力、资源及工具。

从对上述两个目标感兴趣的研究人员的角度出发，本章包含了对这些初始步骤和考

虑的坦率讨论。尽管我们强调基本实验策略的局限性，以及可能遇到的潜在困难，但是大多数研究人员会发现所预期的分析是可行的，并且可能对有机体内转录调控的分子机制获得有价值的见解。

实验策略

新转录因子的表征方法

已经开始把注意力集中于特定转录因子的研究人员通常对鉴定该因子的直接靶基因、阐明该转录因子参与的生物网络和（或）理解该因子调控基因表达的机制感兴趣。他们的首要目标是鉴定其调控区与目的因子直接接触的靶基因，这种接触导致靶基因表达直接的正向或负向调控。大多数转录因子也会通过改变直接靶基因的表达水平，间接调控许多基因的表达，这是因为直接靶基因可能编码调节其他基因表达的蛋白质。

鉴定转录因子的直接靶基因和某些情况下的间接靶基因，是接近第二个目标的第一步。对生物网络的洞察来自于对各种转录因子与在特定生物事件中起作用的信号分子间的相互作用的理解（Blais and Dynlacht 2005）。因此，为了恰当调控，目的转录因子的靶基因通常需要一个或多个其他因子，这种证据往往是阐释生物网络的出发点。另一个共同的出发点是鉴定已知控制生物过程的多个特定因子直接调控的基因。一个新兴的转录网络的例子展示在图 2.1 中，此图来自 Stuart Orkin（Kim et al. 2008）实验室进行的一项研究，显示了胚胎干细胞（ES 细胞）暂定多能性网络。从全基因组染色质免疫沉淀（ChIP）实验（见第 9 章）得出的转录因子与靶基因的联系，旨在鉴定与 ES 细胞

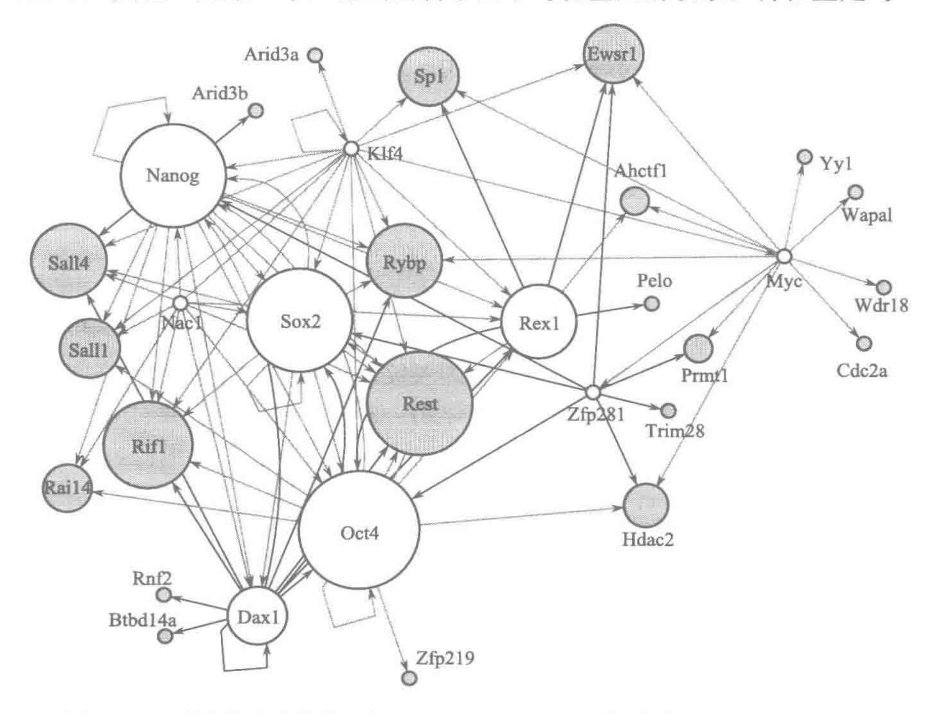


图 2.1 小鼠胚胎干细胞暂定多能性网络（经 Elsevier 许可，转载自 Kim et al. 2008，© 2008）。

多能性有关的 9 种不同转录因子的直接靶基因；在微阵列数据背景中分析转录因子与靶基因的联系，这些微阵列数据检测了暂时鉴定为这些因子直接靶标的基因的表达水平。

第三个目标是弄清楚目的因子调控其直接靶基因的机制，往往同样令人感兴趣和重要。正如第 1 章所述，结合特定 DNA 序列元件的大多数转录因子通过与特定的共激活因子和（或）共抑制因子复合物相互作用行使其功能，这些特定的共激活因子和（或）共抑制因子复合物导致染色质结构的变化、调控转录前起始复合物的组装，或者通过其他策略在转录调控中起作用。转录因子可调控基因表达的许多机制仍然是活性研究的主题，但这些机制可以包括很多作用，如它们在引导靶基因进入特定亚核结构域和促进活性染色质中心组装中所起的作用（见第 1 章、第 5 章和第 9 章）。

鉴定转录因子的靶基因

自发现第一个识别特异性 DNA 序列元件的真核转录因子后的近 20 年中，鉴定转录因子的直接靶基因是一项重大挑战，特别是在哺乳动物和不服从经典遗传学的其他生物中。当获得了新基因的启动子或增强子的 DNA 序列时，通常就揭示了潜在的靶点，表明存在对其偏好识别位点已有了一些认识的因子的可能识别位点。也常用转录因子及其潜在靶基因的表达模式知识来评估转录因子直接调控基因的可能性。

在作出特定因子调控某一基因的假设之后，往往通过证明该因子确实可以在体外结合目的 DNA 元件而获得这一假设的支持。通常通过使用核提取物或重组转录因子的凝胶迁移或者 DNase I 足迹研究，以证明因子与目的 DNA 元件的体外结合。此外，有时有可能表明，假定的识别位点在两个物种（如小鼠和人）之间是保守的，以及可进行瞬时转染实验，以确定在培养细胞中转录因子的过表达是否可调控在目的调控区控制下的报告质粒（有关用于检验转录因子与假定靶基因之间关系的策略的深入讨论，见第 8 章）。从 20 世纪 80 年代晚期至 20 世纪 90 年代初开始，在最佳方案中，通过同源重组靶向破坏编码转录因子的基因，可洞悉转录因子和假定靶基因之间的关系。然而，在许多情况下，这种功能缺失实验是不充分的，因为转录因子家族的不同成员之间的冗余往往会掩盖因子在调控靶基因中的作用；相反，转录因子的缺失可能间接地改变假定靶基因的表达，致使研究者错误地建议一个基因是转录因子的直接靶基因。

在 20 世纪 90 年代晚期和 21 世纪最初的几年中，出现了一些新的技术，这些技术极大地增强了研究人员鉴定已知转录因子及新转录因子直接靶基因的努力。然而，正如下文所述，这些技术都伴随着其自身的局限性，认识到这一点很重要。一项主要的进展是大量生物基因组序列的完成。当与目的因子的首选 DNA 识别序列的知识相结合时，目前全基因组搜索潜在的识别位点是可能的（Elnitski et al. 2006）。由于现在可以得到如此大量的生物的基因组序列信息，这种搜索可限定于在物种间保守性好的潜在识别序列。研究人员也可以将注意力集中在靠近已注释基因 5′端的序列上。通过限制搜索，研究人员可聚焦于那些似乎最有可能在功能上是重要的识别序列。

尽管能够在全基因组里分析目的因子的潜在结合位点是意义非凡的进步，但是这种方法的能力是非常有限的，因为它完全取决于相关知识的深度，即特定蛋白质可识别的 DNA 序列模体，以及关于基因组中识别序列的特定假设。正如第 6 章所讨论的，对于任何因子，还未获得可支持转录因子功能的 DNA 序列模体的完整知识。虽然大量因子

的一致识别序列已经界定，但这些一致序列通常对应于亲和力最高的结合位点。大多数因子可高效地与它们以低于最佳亲和力结合的位点相互作用，并且认为有些因子通过与识别邻近序列的因子的协同结合通过低亲和力位点频繁地行使功能（第 6 章）。相反，控制区内一致序列识别模体的存在并不能保证因子与该模体发生功能性的相互作用，因为其功能的行使可能要求一组合适的其他因子的邻近识别序列。而且，同一家族的多个成员都有可能与一致位点相互作用。

促进发现转录因子直接靶基因的第二个重大进步是 ChIP 分析的发展，以及全基因组 ChIP 策略，如 ChIP 芯片（ChIP-chip）和 ChIP 测序（ChIP-sequencing，ChIP-Seq）（见第 9 章）。如果能获得目的因子的高品质抗体，那么现在这些方法使迅速鉴定活细胞内大量的基因组结合位点成为可能。然而，正如在第 8 章和第 9 章所述，ChIP 策略具有很大的局限性。特别是，越来越多的证据显示，在体内许多因子结合大量的位点并没有明显的功能结果。ChIP 策略也存在技术上的局限性，包括某些因子缺少高品质抗体、抗体之间交叉反应（即使是那些被认为是高特异性的抗体）、获得足够数量的相关细胞类型会遇到的困难，以及准确解释微弱信号的挑战（第 9 章）。

极大地增强鉴定靶基因的努力的第三个进展是 RNA 干扰（RNAi）技术的发展（第 8 章），这项技术往往与基因表达分析和 ChIP 相结合。使用转录因子编码基因的遗传缺失或突变来鉴定转录因子的靶基因已有多年；在典型的研究中，使用基因表达分析[Northern 印迹、反转录-聚合酶链反应（RT-PCR）和现代微阵列分析]鉴定在转录因子缺失的情况下错误表达的基因。然而，对于哺乳动物来说，转录因子编码基因的缺失是费力的，并且通常需要在小鼠 ES 细胞中应用同源重组，然后囊胚注射 ES 细胞产生基因敲除小鼠。已经使用其他策略进行转录因子功能缺失的研究（如显性失活和反义 RNA），但这些方法往往不可靠。

在 21 世纪初，RNAi 成为一种可靠且快速的策略，用于在哺乳动物细胞和来自许多其他真核生物的细胞中进行功能缺失研究（第 8 章）。经过 RNAi 敲低转录因子后，可用微阵列鉴定错误表达的基因。此外，ChIP 结果可与基因表达结果结合，以确定因子是否与候选靶基因的控制区直接联系，或者可以同时分析微阵列结果与 ChIP-chip 或 ChIP-Seq 实验结果，以鉴定一些既在因子缺失的情况下错误表达又与因子直接联系的基因。尽管实验技术的这些组合会特别强大，但是每个技术的局限性往往使结果难以评价。特别是，虽然 ChIP 实验可提供证据，表明因子与在因子缺失情况下错误表达基因的控制区直接相互作用，但是没有其他实验的话，仍然难以得出该因子与其假定靶基因之间联系的有力结论。由于功能缺失实验将鉴定直接和间接两种靶基因，并且 ChIP 实验将揭示功能性和非功能性两种结合位点，所以假命中（false hit）是可以预期的，但假命中的数量则难以预测。

重要的是，虽然基因组序列、微阵列、基于 ChIP 的技术及 RNAi 的可得性极大地增强了鉴定转录因子靶基因的努力，但是这些工具的使用仅仅代表了第一步。最终，广泛的验证是必须的。用于进一步检查和验证转录因子与假定靶基因之间关系的策略将在第 8 章中描述。

阐明作用的机制

除了鉴定和验证转录因子的靶基因，许多研究人员对弄清楚目的因子调控基因表达的机制感兴趣。当试图阐明调控机制时，首要目标通常是确定该因子被招募到 DNA 靶点的机制，以及与该因子相互作用而使因子发挥其生物学功能的蛋白质。对于序列特异性 DNA 结合蛋白，DNA 招募的理解通常涉及蛋白质的首选 DNA 识别序列的划定。通过使用结合位点选择分析及最近的基于阵列的方法，可确定共有识别序列（见第 11 章）。对于被认为靶向特定基因座但不以序列特异性方式与 DNA 结合的蛋白质（如共激活因子和共抑制因子），必须进行蛋白质-蛋白质相互作用的研究。鉴定相互作用的蛋白质对于理解转录因子或共调控因子的作用机制这一目标也是至关重要的。

可用数种实验方法来鉴定能够与目的因子相互作用的蛋白质。在第 10 章中对这些方法进行了简要地讨论。然而，如上所述，主要的挑战是确定所检测到的相互作用的生理相关性。在理想情况下，目的蛋白将是一个高度稳定的多蛋白复合物的一种组分。如果是这样，可通过柱层析（最常见的是免疫亲和层析）将该复合物纯化到同质，并通过质谱鉴定结合在一起的蛋白质（见第 7 章和第 10 章）。如果该复合物是高度稳定的，并且在缺乏蛋白质过表达的情况下也能纯化，以及如果该复合物的组分以近乎化学当量的水平出现，那么可以相当自信地断定所发现的相互作用是生理相关的，并可能对生物学功能及目的蛋白的作用机制是重要的。基于相互作用蛋白的身份和属性，可设计实验进一步分析作用机制。例如，相互作用蛋白可能具有酶活性，这种酶活性催化组蛋白或参与转录调控的其他蛋白质的翻译后修饰。

虽然许多参与转录的蛋白质组装成高度稳定的多蛋白复合物，如 RNA 聚合酶、TFIID 和调解因子，但大多数典型的 DNA 结合转录因子并非如此。由于这个原因，确定典型因子调控转录的机制会非常困难。当目的蛋白与其他蛋白质不稳定结合时，一种通常用于鉴定相互作用蛋白质的实验技术是酵母双杂交筛选（见第 10 章）。虽然在许多场合下该技术取得相当大的成功，但是难以确定在一个典型的筛选中鉴定的许多相互作用蛋白质中，哪些蛋白质真正与内源性转录因子的功能是相关的。第二种实验技术是以正常表达目的蛋白的细胞进行免疫沉淀分析。然后可用 Western 印迹（Western blot）显示候选相互作用蛋白质是否存在于免疫沉淀物中。可将这种方法与酵母双杂交筛选结合起来，作为一种测验，用于澄清在自然表达目的蛋白的细胞中是否也能检测到在酵母中检测到的相互作用。

如第 10 章中所述，可用其他方法鉴定与目的因子相关联的蛋白质，这些方法包括各种串联亲和纯化（TAP-标签）方法及各种柱层析方法。串联亲和纯化方法有时可用于分离细胞提取物中不稳定的多蛋白复合物。在柱层析方法中，可以使核提取物通过含有固定化转录因子的树脂。如上所述，虽然所有这些策略都能提供有价值的信息，但是研究结果一般都没有定论，除非蛋白质形成可易于纯化的高度稳定的复合物。如果蛋白质微弱地相互作用，或者如果每种蛋白质只有一小部分装配成稳定的复合物，那么需要其他的实验来检验该相互作用真正的生理相关性。之所以这样说，是因为许多蛋白质的一小部分可能在粗提物中与许多其他蛋白质任意地相互作用，这种相互作用没有任何生理相关性。因此，除非该相互作用是高效的且高度稳定，否则必须进行功能性研究，最

好用生理相关性分析，以检验相互作用的意义。

有数种功能分析方法可用于测验蛋白质-蛋白质相互作用的功能相关性。然而，这些分析方法中的许多分析方法超出了本书的范围，依赖于蛋白质的过表达，这会产生误导性的结果，因为细胞内蛋白质浓度的增加可以促进蛋白质-蛋白质的相互作用，但当蛋白质在生理浓度下存在时，这种相互作用不会发生。一种替代性的首选策略是通过RNAi或其他功能缺失方法消除相互作用蛋白质的功能，以确定该蛋白质的缺失是否消除或改变目的转录因子的功能（即该蛋白质调控靶基因的能力）。虽然依靠RNAi的功能缺失研究快速且简单，但是所有功能缺失方法都受制于无力区分直接和间接效应。因此，虽然相对于过表达实验，一般首选功能缺失研究，但其结果必须谨慎解释。

如第10章中所述，为了进一步确信蛋白质-蛋白质相互作用是生理相关的，可进行结构域交换实验。此外，结构域交换实验有助于周密定位目的转录因子中支持该相互作用的残基，并且有助于确定这些残基对该因子的功能是否重要。最好是在可使突变蛋白以生理浓度表达的分析方法中进行功能研究，以避免过表达人造物。虽然这样的分析结果能使人更加确信相互作用是重要的，但是结果必须再次谨慎对待，因为相同的残基可支持与多种不同蛋白质的相互作用。核激素受体为这一点提供了一个杰出的例子，因为研究得非常透彻的受体的相同残基支持与多个共激活因子的相互作用（Lonard and O'Malley 2006；O'Malley et al. 2008）。尽管许多已报道的相互作用的生理相关性已经获得了相当多的实验支持，但是这些受体激活转录的精确机制还没有被定论性地确立。

上述讨论强调了在目的转录因子的作用机制中，初步结果与定论性认识之间的不同。可以在酵母双杂交筛选或免疫共沉淀分析中容易地鉴定潜在的相互作用。虽然过表达或RNAi实验可提供一些相互作用是生理相关的证据，但是需要大量额外的努力以产生令人信服的结论。此外，我们只讨论了一项机制分析中的第一步，它涉及DNA识别位点和相互作用蛋白质的鉴定。在已经达到这些初始目标后，需要更多的时间和精力来确定蛋白质-DNA和蛋白质-蛋白质相互作用调控基因表达的机制。在许多情况下，最终需要第12章和第13章中描述的那些体外方法。因此，在着手对新转录因子进行机制分析之前，重要的是要考虑对科研文献做出真正有意义的贡献所需要的资源和付出的努力。

分析新基因的调控

初始步骤

现在，我们重点讨论一下对分析新基因的调控感兴趣的研究人员的初步考虑。图2.2是初始转录调控分析的概念和技术进展的流程图。在开始分析之前，一个非常重要的步骤是确定调控的潜在方式是否真的在基因水平（图2.2）。或许仅仅知道某种蛋白质的生化活性在一种细胞类型中是可检测到的，但在另一种细胞类型中是检测不到的。这可能归因于蛋白质活性及其稳态丰度的调节，或者蛋白质mRNA稳态丰度的调节。因此，下述三种情况是可能的。

（1）如果蛋白质的活性和丰度，以及其mRNA丰度在两种细胞类型中不同，则活性至少部分在mRNA（即基因表达）水平上被调节。

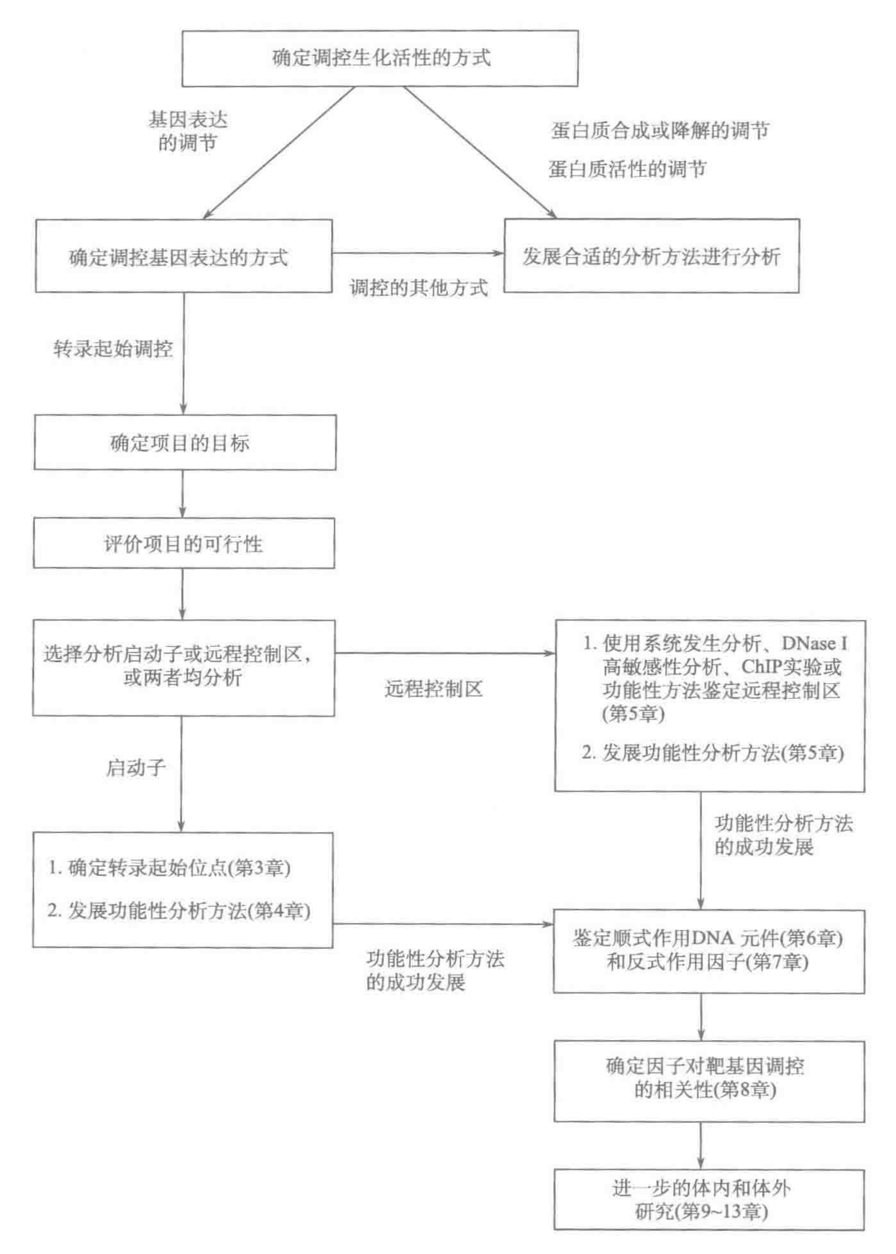


图 2.2 转录调控机制的分析。

(2) 如果蛋白质的活性和稳态丰度（非 mRNA 的丰度）在两种细胞类型中不同，则活性主要在蛋白质水平上被调节，最有可能通过差异翻译或蛋白质降解调节。

(3) 如果蛋白质的活性（而非蛋白质丰度或 mRNA 丰度）在两种细胞类型中不同，则活性最有可能通过翻译后修饰或者与其他蛋白质或配体的物理相互作用调节。

为了区别以上三种情况，需要在两种细胞类型中简单地比较蛋白质的生化活性、稳态蛋白质数量及稳态 mRNA 数量。通过多种方法可测量稳态蛋白质的水平，包括标准

的免疫印迹或免疫沉淀分析。通过 Northern 印迹、定量 RT-PCR，或者任何监测稳态 mRNA 水平的其他方法可确定稳态 mRNA 的水平［如引物延伸、RNase 保护或 S1 核酸酶分析（见第 3 章）］。

此处，我们专门讨论一下第一种情况。在这个情况中，调控主要在稳态 mRNA 丰度水平上。尽管如此，数量巨大的蛋白质的生化活性受到直接调控，并且大量的蛋白质似乎受到翻译或降解的调控。如果蛋白质的活性直接受到调控，那么可开展蛋白质的生化和结构研究，以鉴定相关的翻译后修饰、相互作用蛋白质，或者其他相互作用分子。如果调控是在稳态蛋白质丰度水平上，那么可开展研究，以区分翻译调控和蛋白质降解，然后通过实验来剖析相关的途径。

如果调控主要是在稳态 mRNA 丰度水平上，则重要的是要牢记 mRNA 的调控不一定涉及转录的调节。在稳态 mRNA 丰度中的差异也能在多个水平上调控，包括 mRNA 差异降解、mRNA 转移或者前体 mRNA 剪接。

为了确定调控是否在转录水平上，一般使用 4 种分析方法：核连缀分析（专题 2.1 和方案 2.1）、通过实时 RT-PCR 分析前体转录本、ChIP 技术监测 RNA 聚合酶结合或基因座其他特性，以及转染分析中的启动子-报告基因活性分析。正如在专题 2.1 和方案 2.1 中所描述的，核连缀是证明基因在转录起始水平上受调控的经典方法。然而，因为其难度和对相对大量的放射性要求，这种方法已不再普遍使用。较新的方法取得同样有说服力或更具说服力的结果，但花的精力却更少。

新兴的策略之一是利用 RT-PCR 或实时 RT-PCR 监测基因在其不同表达状态中的前体转录本的水平（Lipson and Beserga 1989；Reddy et al. 1994；Su et al. 2004）。这通常是通过使用退火至内含子的 RT-PCR 引物和退火至相邻外显子的反向引物来完成的。这些引物将允许前体转录本的扩增，但它们不会扩增成熟的 mRNA 或被切除的内含子。现阶段认为前体转录本的水平反映转录水平比反映 mRNA 水平更加准确，因为据认为通过 mRNA 剪接，前体转录本总是被快速加工。因此，在不同表达状态中的前体转录本水平的变化通常反映转录中的变化，而不是 RNA 稳定性中的变化。这种方法的一个警告是，一些基因的前体转录本将受到差异稳定性或前 mRNA 剪接差异速率的调节仍然是可能的。值得注意的是，最近已描述了在全基因组水平上监测新生转录本的策略（Cheadle et al. 2005；Ohtsu et al. 2008）。

现在，ChIP 是另一种常用于证明基因在转录水平上被调控的分析方法。RNA 聚合酶 II、组蛋白修饰及关键调控因子结合 ChIP 信号中的差异可提供转录调节负责特定基因调控的强有力的证据（Sandoval et al. 2004）。这种方法需要注意的是，在有预结合 RNA 聚合酶和存在一般与活跃基因有联系的组蛋白修饰的情况下，有可能激活一些不活跃的基因。对于这些基因，使用 ChIP 来监测转录单位 3′ 端附近，而不是启动子附近的 RNA 聚合酶水平，可提供转录调控的证据。

有时用于获得目的基因转录调控证据的最后一个方法是使用启动子-报告基因质粒的转染分析。如果报告基因活性反映了内源性基因的 mRNA 水平，这一结果也许提示启动子活性的调控负责或至少部分地负责 mRNA 丰度的调控。该策略需要注意的是，它只能应用于容易转染的细胞。第二个需要注意的是，有时在启动子-报告基因分析中

转录未被正确地调控，即使转录的调节可能完全负责内源性基因的调控。

虽然严格地确定目的基因的 mRNA 丰度是否在转录水平被调控是重要的，但是转录似乎是调控大多数基因的主要方式，因为选择性前 mRNA 剪接在调控来自大量基因的选择性 mRNA 剪接产物的表达中起到关键作用。下面的讨论突出分析控制区（如启动子和增强子）和转录调控因子前考虑的问题。

专题 2.1 核连缀转录分析

核连缀技术（图 2.3）（Derman et al. 1981；Greenberg and Ziff 1984；Powell et al. 1984；Ucker and Yamamoto 1984）一般在不同水平表达目的基因的细胞群中进行。例如，当研究可诱导基因时，需要使用诱导前和诱导后的细胞群。对于组织特异性基因，对表达目的基因的细胞与缺乏目的基因表达的细胞进行比较。

该程序中的第一步是在冰上裂解细胞膜而不破坏核膜（图 2.3）。在裂解和冷却时，正在延伸的聚合酶分子"冻结"在正在活跃转录的基因上。然后，通过离心沉淀细胞核

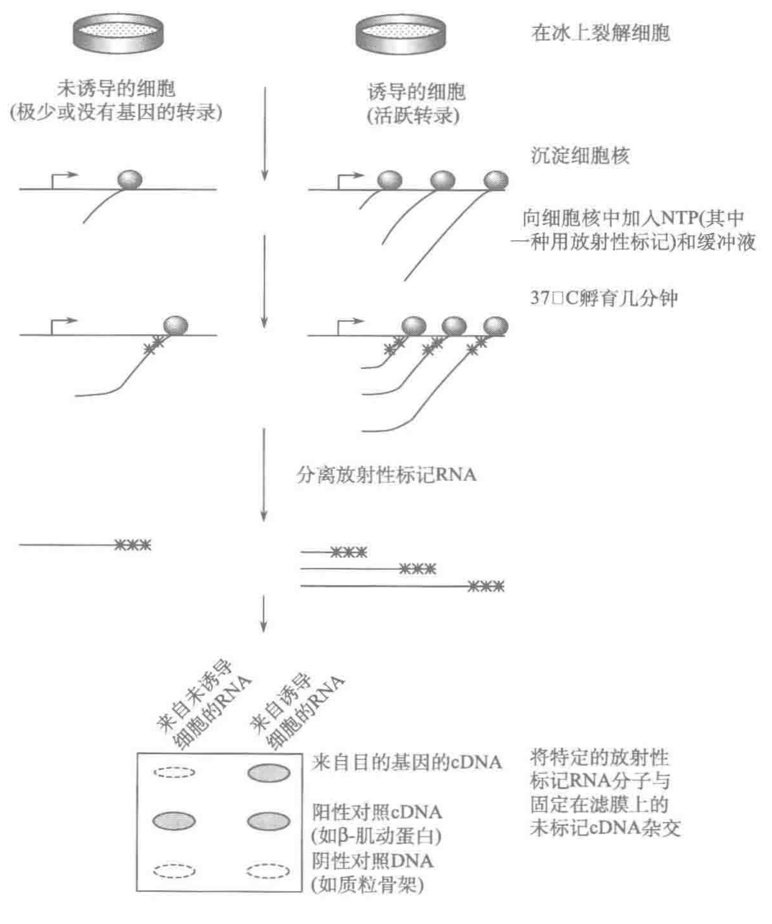


图 2.3 核连缀分析。

并洗涤。之后，将产生的细胞核悬浮在一种反应缓冲液中，该缓冲液含有^{32}P标记的三磷酸核苷（NTP）和三种未标记NTP。将混合物在37℃孵育很短的时间（5～15 min），让已与DNA结合的聚合酶分子恢复延伸。在分离的细胞核中，从头转录起始不会发生。分离放射性标记的前mRNA，并与含有来自目的基因的固定化未标记DNA序列的膜进行杂交。通过放射自显影和密度仪或磷屏成像仪分析定量特定的放射性标记RNA。

关键的理论概念是，当裂解和冷却细胞时，通过放射自显影观察到的特定放射性标记RNA的量与该基因结合的聚合酶分子数大约成比例。相应地，聚合酶分子数，有时称为“聚合酶密度”，应该与从该基因转录起始的速率成比例。因此，如果基因在诱导前没有转录，而在诱导后活跃地转录，则诱导后特定放射性标记RNA的量将比诱导前更高，从而反映出稳态mRNA丰度的变化。相比之下，如果转录起始速率在诱导后不改变，那么特定放射性标记RNA的量将在诱导前后相似。如果稳态mRNA的丰度在诱导期间改变，这一后者核连缀结果可能提示mRNA丰度受到不同机制的调控，如mRNA的稳定性。

考虑达到明确目标所需的时间投入和资源

在开始转录调控的分析之前，必须要考虑的两个问题是所需的工作量和资源。由于分析的起始步骤较易完成，故往往低估了工作量和资源。

初步而有限地洞察转录调控机制的两种通用策略

对目的基因转录调控全面而基本的分析，通常从一个或者多个重要控制区的鉴定开始，如启动子、增强子、沉默子或者是基因座控制区（LCR）（见第1章）。正如第3章中所描述的，启动子鉴定一般要求转录起始位点的准确定位。如第5章中所述，可采用多种方法鉴定增强子、沉默子以及LCR（统称为远程控制区），这些方法包括来自多个物种基因组序列的系统发生分析、全基因组组蛋白修饰分析，以及DNase I高敏感性方法。鉴定控制区后要进行：①发展检查启动子或远程控制区功能的分析方法（第4章和第5章）；②划定控制区内重要的DNA序列模体（第6章）；③鉴定与模体相互作用的因子（第7章）；④证明该因子真正调控该基因（第8章）。随后实验也是相当普遍的，这些实验以一组确定的转录因子检查在自然染色体环境中（第9章）或者在无细胞转录分析中（第12章和第13章）的调控。

对参与基因调控的因子获得有意义的理解所需要的时间和工作量是巨大的。以更快速分析的局限性为例，我们注意到，对研究人员来说，获得新基因的启动子序列，并使用数据库（如TRANSFAC或JASPAR）鉴定类似于已知蛋白质一致识别序列的DNA模体是相当普遍的（见第6章）。如果鉴定到潜在的控制元件，则可以通过在标准瞬时转染分析中测验特定置换突变评价这些潜在控制元件的相关性（见第4章和第6章）。如果突变降低了启动子活性，那么可以通过使用重组蛋白或核粗提物进行凝胶迁移率变动分析（EMSA），确证预测的蛋白质在体外结合这些控制元件的能力（见第7章和第11章）。然后，研究人员可用ChIP分析来确定预测的因子是否在体内与控制区结合。

在许多情况下，这种方法已经提供了对基因调控机制的洞察，但对于大多数情况

下，必须认为获得的信息是初步的。在缺少无偏突变体分析以鉴定重要控制元件的情况下，难以确定通过数据库鉴定的 DNA 是否发挥主要作用或次要作用，当然这是相对于一些未鉴定的 DNA 模体在被调控的启动子活性中发挥的作用而言的。而且，已发现结合到该模体上的蛋白质通常是一个 DNA 结合蛋白大家族的成员，该家族的成员具有相似的位点识别特异性。需要投入相当大的工作量来明确地确定哪个家族成员负责目的基因的调控（正如本章前面所述，仅有在 ChIP 分析中的结合和 RNA 干扰实验中的表达缺失是不够的）。该基本方法的最后一个局限性是，它强烈地偏向于先前已表征的而非未表征的转录因子。总的来说，这种策略提供了有效的出发点，但需要更多的工作以严格确定基因是如何被调控的。

快速而初步地洞悉基因调控机制的第二个策略涉及对启动子或远程控制区发展恰当的功能性分析方法，其中最简单的是用报告质粒进行瞬时转染分析（第 4 章）。在发展功能性分析方法之后，研究人员可进行控制区的系统缺失分析（第 6 章）。缺失分析可能确定含有重要控制元件的 DNA 短片段（如 50 bp），而 EMSA 或 DNase I 足迹法（第 7 章和第 11 章）可用于鉴定与该区域相结合的蛋白质。这种策略比基于计算机的策略偏向性更少，并且有更大的潜力鉴定到先前未表征的 DNA 结合蛋白。尽管如此，如果通过 EMSA 或 DNase I 足迹法鉴定的蛋白质似乎是新的，则该蛋白质的检测对于其表征、鉴定及其在目的基因调控中作用的验证只是一个起点。

确定项目目标

上面的讨论把注意力集中在确定项目目标的重要性，以及弄清楚达到这些目标所需的时间投入和资源上。工作量和资源随目标的不同而有极大的差异。在一个极端，为了在转基因小鼠特定组织中表达蛋白质，初始目标可能只是鉴定能在该组织中指导转录的控制区。这一目标往往会（但并不总是）涉及一系列相对简单的实验，包括鉴定转录控制区及在转基因小鼠分析中检验功能活性，或者在转染分析中进行初步的功能研究。另一个目标可能是确定已知的转录因子是否是目的基因的直接调节者。虽然难以获得确凿的证据（见上文和第 8 章），但所需的工作量依然相对适中。然而，在另一极端，目标可能是定义和鉴定新的转录因子，并严格地确证它们的确是目的基因的重要调控者。尽管技术正以迅猛的速度在改进，但实现这些目标将需要几年，更可能是好几年的时间。在确定调控因子之后，如果目标是阐明各种因子相互一致行动以调控通过 RNA 聚合酶 II 的转录起始的其详细机制，则可能需要整个职业生涯，因为对于任何真核基因尚未达到此目标。

应该强调的是，在综合分析的过程中，可以取得许多重要的进展。例如，控制元件和控制区潜在调控蛋白的基本划定代表重要的一步。其他重要的进展还有：①先前未表征的 DNA 结合蛋白的鉴定和表征；②严格证实 DNA 结合蛋白对基因调控相关性的实验；③洞悉已鉴定 DNA 结合蛋白生物学功能的基因破坏实验；④调控 DNA 结合蛋白活性的分子机制和信号转导途径的分析。因此，尽管难以达到对基因调控机制的完全了解，但是追求这一目标可能会导致许多极其重要的进展。

评估分析的可行性

基于可用的工具，第二个需要考虑的问题是分析的可行性。在最重要的问题中包括用于功能和生化研究的细胞的来源，以及成功地发展恰当的功能分析方法。

用于功能性研究的合适细胞来源

在综合分析的初始阶段中，主要的目标是在启动子或远程控制区中划定关键的控制元件，通过比较野生型控制区和突变型控制区的功能可鉴定这些元件。由于对染色体中内源基因进行系统性突变分析一般是不合理的，因此必须使用人工分析方法（见第 4 章）。对这样的分析最有用的分析方法涉及细胞瞬时转染或稳定转染，在这些转染实验中使用含有野生型或突变型控制区的质粒，其中控制区调控一个合适的报告基因的转录。这些类型的分析统称为功能性分析，因为它们提供对控制区调控功能的半定量测量。

为了确定研究特定基因的可行性，关键的考虑之一是能否得到用于功能性分析的合适的细胞来源。对于哺乳动物基因来说，这样的分析通常依赖于特定细胞系或细胞群的获得，这类细胞能在培养基中生长并能在转染中吸收 DNA。在大多数情况下，能表达目的基因的转化（永生）细胞系是最合适的。被转化的细胞系一般易于在培养基上生长，并能保持足够的数量，而且，它们通常比原代细胞更适合转染。

正如第 4 章所述，使用转化细胞系的缺点是调控目的基因的转录因子可能相对于原代细胞有所改变，在原代细胞中调控基因的因子可能在转化细胞中不表达。此外，还有许多例子说明转录因子在转化细胞中表达，而在同种细胞类型的非转化细胞中却不表达。尽管如此，如果目的基因在转化细胞中表达，应该有可能确定在该细胞系中负责其转录的因子。在阐释了这些因子和调控机制之后，可以设计在原代细胞或动物中的随后实验，以评估所鉴定的因子的重要性和所提出的调控机制的有效性。例如，如果在转化细胞中的研究导致了这样的假设，即靶基因的表达需要特定的转录因子，那么可通过 RNAi 或同源重组扰乱编码转录因子的基因，以确定该因子对于原代细胞中靶基因表达是否的确必不可少。此外，对于此分析，如果可以获得足够数目的原代细胞，那么在原代细胞中进行 ChIP 实验是可能的（见第 8 章）。

如果目标是确定在特定发育阶段或对特定外源刺激作出反应期间基因被诱导或抑制的机制，那么在分析的初始阶段，使用可以模拟诱导或抑制事件的细胞系将是有益的。可比较诱导或抑制前后控制区的活性，来确定相关的 DNA 元件和转录因子。一种不太吸引人但有效的替代方法是比较表达和不表达目的基因的转化细胞系。例如，为了研究负责免疫球蛋白（Ig）μ 基因细胞特异性活化的机制，对 B 淋巴细胞和普通的非淋巴细胞系中（Ig）μ 增强子的活性进行了比较（Banerji et al. 1983；Gilles et al. 1983）。这种方法没有诱导系统理想，原因是不同细胞系可能具有很多不同的特性，这些特性可能产生误导性的结果。尽管如此，对于研究细胞特异性基因调控来说，该方法是必不可少且有价值的。

如果目标是对新基因进行全面分析，那么缺乏合适的永生细胞系会带来严重障碍。在有些情况下，如果可以发展出一个有效的转染或反转录病毒转导程序，那么就可以在

培养基中至少可以生长几天的原代细胞中进行研究。第二个替代方法是在转基因鼠中进行功能研究。转基因鼠对鉴定调控区是有用的，并且可用于有限的突变分析，以鉴定不相关联的控制元件（第 5 章）。然而，由于高成本和工作量的原因，这一替代方法对于启动子或增强子的深入分析一般是不可行的，最终的解决办法就是利用体外转录分析。如果可以获得表达目的基因的丰富的细胞来源，这一分析方法有时是可行的。这些替代方法将在第 4 章和第 12 章中详细地讨论。

总而言之，基因调控分析的一个重要的必要条件是，为顺式作用控制元件的功能剖析得到合适的细胞来源。如果可以得到表达内源目的基因的转化细胞系，而且可以通过标准技术转染，那么有可能对调控机制获得相当大的认识。如果不能得到合适的细胞系，那么需要考虑其他更具挑战性的替代方法。

用于蛋白质抽提物制备的细胞来源

在已经确定了负责调控某一基因的 DNA 序列元件之后，必须鉴定并表征与关键元件相互作用的蛋白质。用于功能研究的细胞系通常对于制备用于生化分析的无细胞提取物是有用的。转化细胞系的主要优点是，它们通常是同质的、能够大量生长，以及适于在功能研究和生化研究之间进行直接比较（如果同一细胞系用于两种目的）。转化细胞系的缺点是，目前的这些因子可能未准确代表原代细胞中存在的那些因子。

前面一节中还未讨论的一个问题是可获得的细胞的数量。这个问题对于生化分析比对于功能分析往往更重要，特别是如果必须纯化蛋白质时（见第 7 章和第 12 章）。如果细胞是非贴壁的，则转化细胞对于这一目的是有用的。非贴壁细胞往往可以在飞旋培养物中大量生长（通常在特殊设计的带有搅拌棒或悬浮搅拌臂的细胞培养瓶中），或者在旋转瓶中（瓶中 90%充满培养液）生长。对蛋白质纯化而言难以获得足够数量的贴壁细胞，但在有些情况下，可使贴壁细胞适应非贴壁生长。尽管转化细胞经常用于蛋白质纯化，但也能考虑使用原代细胞。利用动物组织纯化和表征转录因子已获得很多成功。NF-κB 提供了转录因子的一个重要的例子，该转录因子就是从一种主要组织，特别是兔肺细胞中纯化并用于多肽序列分析的（Gosh et al. 1990）。如果能鉴定转录因子的组织来源，动物组织可能提供廉价且耗时更少的用于蛋白质纯化的细胞准备方法。然而，仔细地表征提取物是必需的，因为有些组织对蛋白质降解异常敏感。正是由于这个原因，进行比较各种新鲜和冷冻组织的试验性研究是重要的。

成功开展一个适当的功能分析

如果可得到用于蛋白质和功能研究的合适的细胞来源，那么通常将可能取得解析调控机制的重大进步，然而，项目的成功主要取决于一种切实可行的启动子或远程控制区功能分析方法的发展。在分析方法的发展中已经遇到一些困难，这些困难与用于转染实验的合适的细胞系的获得无关。

对于某些启动子来说，由于仍然费解的原因，一直非常难以用标准的瞬时转染分析检测启动子活性。例如，末端转移酶（Dntt）启动子在有效表达内源基因的前 B 细胞系和前 T 细胞系中的瞬时转染分析中功能不佳（Lo et al. 1991）。启动子可能只是在缺乏远程增强子时较弱，或者启动子的未知特性可能阻止其在人工分析方法中行使功能。在其他情况下，启动子活性易于检测，但并不显示出内源基因的细胞特异性。目前还不

清楚具有这种表型的启动子在它们各自基因的细胞特异性中是否不起作用，或者细胞特异性因子在人工分析方法中是否未正确行使功能。例如，改变的染色质结构及大量拷贝数的瞬时转染质粒可能会引起异常的功能（见第 4 章）。

如果主要目标是鉴定和表征特定基因的远距离控制元件（如增强子、沉默子或 LCR，其中任何一个距离启动子都相当远），则研究的可行性可能受到研究人员鉴定相关区域的能力的限制。已经在内含子或基因上游或下游几万碱基对内鉴定出控制区。第 5 章中所描述的策略可能导致远程区域的鉴定，但在有些情况下这些方法将会失效。

有时在开始分析之前可以评估遇到这些类型困难的可能性，因为已经证明有些类型的基因比其他类型的基因更适合进行转录分析。例如，很多快速诱导基因，包括早期应答和可诱导细胞因子基因（如 c-*fos* 和白细胞介素-2［IL-2］基因；Treisman 1985；Siebenlist et al. 1986），更适合进行转录分析，因为这些基因的启动子对转录的快速诱导起关键作用，并且在瞬时转染分析时通常可以检测到该类基因的可诱导启动子活性。相比之下，一些发育调控基因（如 *Dntt* 和 *Rag* 基因）的启动子在标准瞬时转染分析中功能不佳，并且有时不能反映内源基因的恰当调控。因此，通过纵览文献确定研究相似类型基因的其他研究者取得成功的水平，可了解自己的研究成功的可能性。

总之，综合性转录调控分析的可行性取决于用于功能研究和蛋白质研究的合适的细胞来源的获得，以及为目的控制区分析成功发展一种合适的功能性分析方法。在进行实验之前，可对项目的可行性进行初步评估。如果可得到合适的细胞来源，那么就可以很容易地启动这个项目，但是在进行了发展一种功能性分析方法的初步尝试之后，需对该项目进行重新评估。如果目的控制区未显示出可检测的或者在标准的瞬时转染分析中正确的调控活性，则可以尝试更复杂的分析方法（例如，稳定转染或转基因分析），或者可能需要重新考虑该项目的目标。例如，在缺乏功能性分析方法的情况下，通过使用依据计算评估预测与控制区结合的因子的抗体进行 ChIP 研究，研究人员对可能调控控制区的因子就可能能够获得不太全面和不太确凿的认识。另外，如果在功能分析中观察到了预期的活性，则应该有可能阐明基因的主要调控机制。

启动对新基因的综合转录调控分析

从启动子或远程控制区开始

在证明基因是在转录起始水平上调控且目标合理可行之后，必须作出关键决定，即是开始启动子区的分析，还是开始远程控制元件（如增强子、沉默子和 LCR）的全面搜索，第三种选择是同时开始两种研究。开始启动子分析的策略在第 3 章和第 4 章中描述，远程控制区的研究在第 5 章中描述。尽管调控机制的全面理解将需要理解每个控制区，但是从以上两个方向着手实施该项目所必需的资源可能是不可得到的。从分析启动子开始，还是从分析远程调控区开始，或者从分析两者同时开始，其决策必须基于多个标准，而其中最重要的是研究者的兴趣。

对于大多数基因来说，重要的调控事件将在启动子和远程控制区都发生是很有可能的。例如，Ig μ 基因的启动子和内含子性增强子（intronic enhancer）对 B 细胞特异性表达都起作用（Calame and Ghosh 1995；Ernst and Smale 1995）。有些研究人员认为研

究远程控制区具有更高的优先权，因为这些区域通常是强大的并且可能控制整个基因座的可及性——远程控制区可能改变整个基因座的染色质结构，从而使转录因子结合到启动子上成为可能。因此，在调控等级上，认为远程控制区可能处于更高的位置。专注于远程控制区的缺点之一是可能难以发现这些区域。研究某些远程控制区（如 LCR）的第二个缺点是，这些远程控制区在标准瞬时转染分析中经常不起作用；相反，需要难度更大、耗时更长的分析方法，如稳定转染或转基因分析（见第 5 章）。

研究启动子的主要优势是启动子位于转录起始位点的附近，根据定义，它紧邻转录起始位点。因此，通过转录起始位点作图，可以明确并相对容易地定位启动子（第 3 章）。此外，在体外分析中，极有可能重现启动子活性的生化机制，这可在分析的高级阶段促进详细的机制研究（第 12 章）。以启动子区域的分析开始转录调控研究的主要缺点是，在调控等级中，它可能位于远程控制区的下游。然而，重要的是要注意，启动子和远程控制区组装成活性染色质中心（ACH），提示在内源性基因座上它们的功能可能是完全相互依存的（de Laat and Grosveld 2003）。

如果有足够的时间和资源，那么最理想的策略可能是同时开展启动子和远程控制区的研究。在完成每种分析的最初步骤之后，选择一个还是两个方向进行研究可能会变得更清楚。例如，在开展的功能分析中，启动子可能模拟或至少部分模拟内源性基因的调控，而分离远程控制区的尝试也许会失败。这种情况可能导致一种有效决策，即研究启动子以确定引起恰当调控的一些事件。相比之下，最初的研究可能导致一种强大的且适当调控的增强子或沉默子的鉴定，但最初的启动子分析可能揭示启动子活性在所用的功能分析中是普遍存在的。在这种情况下，有效的决策可能是专注于增强子或沉默子的调控性贡献，以及评估启动子在后期的贡献。

用于在启动子和远程控制区中进行选择的上述策略对很多基因都是有用的，但对某些特定类型的基因来说，出发点可能会更明显。正如上面所提到的，在细胞受刺激后，被快速诱导的基因，一般如免疫系统中的细胞因子基因以及早期应答基因，似乎将诱导的主要决定物集中到启动子上。这些基因可能有对调控起作用的远程控制区，但是远程控制区可能主要负责调控在特定细胞类型中该基因座的可及性，以便通过蛋白质结合启动子使该基因有望快速激活。因此，对这类基因来说，为了阐明转录的细胞类型特异性，需要分析远程控制区，但是为了阐明快速诱导的机制，启动子通常必须是主要焦点。

总之，为了确定蛋白质活性是否真正在基因表达水平被调控，以及基因表达是否在转录起始水平被调控，转录调控的综合分析必须从实验开始。如果这些实验证明转录起始是重要的调控方式，那么应该在评估项目目标时，仔细地考虑达到这一目标所需的时间投入和资源。评估项目的可行性时也必须考虑可用的工具（如细胞系）。此外，必须决定是对启动子还是远程控制区开始分析，还是对两者同时开始分析。然后，必须分离控制区，并为控制区的分析发展一种功能性分析方法。有了恰当的功能分析方法，取得阐明转录调控机制的重要进展是指日可待的。

尽管相当详尽地研究真核基因调控机制已有 20 多年，但是很多基本问题仍有待解决。对于阐明控制特定细胞类型发育和功能的调控线路或者促成疾病发病机制的异常线路，个别实验室可取得重要进展。或许更引人入胜的是在很多实验室进行的共同研究的

可能性，如果这种可能性变成现实，构成细胞或生物体内基因调控线路基础的整体逻辑将开始出现。

技术

方案 2.1　核连缀分析

核连缀分析被发展为一种证实转录起始速率有助于哺乳动物基因的被调控表达的方法（Derman et al. 1981；Greenberg and Ziff 1984；Powell et al. 1984；Ucker and Yamamoto 1984）。核连缀分析监测基因表达与大多数其他分析（如 Northern 印迹、引物延伸、核酸酶 S1、RNA 保护，或 mRNA 的 RT-PCR 分析）监测基因表达之间的差异是，核连缀分析提供了转录起始频率的测量，而且核连缀分析在很大程度上不受 RNA 稳定性的影响（见本章正文及文献 Derman et al. 1981）。相比之下，其他分析测量的是稳态 mRNA 丰度，唯一的例外是使用 RT-PCR 或实时 RT-PCR 来监测前体转录本的丰度（见正文）。核连缀分析也可用于确定聚合酶暂停或衰减是否有助于基因调控。

简单地说，核连缀分析从含有不同稳态量的 mRNA 或目的蛋白的细胞样品开始（见专题 2.1）。冷却细胞并通透化或裂解细胞膜，这些步骤会导致聚合酶暂停。然后在 NTP 和放射性标记尿苷 5′-三磷酸（UTP）存在的情况下将细胞核在 37℃孵育一小段时间。在孵育期间，新的转录本并未起始，但是当第一次冷却和裂解细胞时，放射性标记核苷酸掺入正在合成的转录本中。在冷却的时候，据认为基因上新生转录本的数量与转录起始频率成正比。为了确定每个样本中新生转录本的相对数量，纯化放射性标记的 RNA 并与膜进行杂交，其中膜上固定着来自目的基因的 DNA。与膜杂交的放射性剂量与新生转录本的数目大约成正比。

已经发展了核连缀方案的多重变种，下面的方案是改编自 Greenberg 和 Ziff（1984）及 Linial 等的方案（1985）。至于附带其他考虑的类似方案，见 Sambrook 和 Russell（2001，p. 17.23）及 Ausubel 等（1994，单元 4.10）。

时线与安排

在开始核连缀程序之前，必须准备好未标记的 DNA，将其固定到与放射性标记的 RNA 杂交的膜上。准备 DNA 时应当考虑以下几个方面。

1. 对于基本的核连缀程序，来自目的基因的双链 DNA 和单链 DNA 都能使用。更长的基因组 DNA 片段将比较短的片段产生更强的信号，但是较短的片段对于监测与基因座特定区域相关联的新生转录本的数量来说是可取的。

2. 为了增加机会，使该分析可以测量不受聚合酶暂停或衰减影响的转录起始的频率，首选来自转录单位 5′端的片段。

3. 与新生转录本互补的单链 DNA 片段优于双链片段，因为反义转录（这在哺乳动物细胞惊人地常见）造成的背景信号将会降低。通过把目的基因片段插入 M13 或噬菌粒载体，可以生成单链 DNA 分子（Sambrook and Russell 2001，p. 3.42；Ausubel et

al. 1994，单元 1.14 与 1.15)。

4. 除了与目的基因对应的 DNA 之外，还强烈推荐阳性和阴性对照 DNA。阴性对照通常是缺乏基因插入的质粒载体。合适的阳性对照将是来自组成型表达基因［如β-肌动蛋白或甘油醛-3-磷酸脱氢酶（GAPD)］的 DNA 片段。

有了纯化的单链或双链 DNA，就可以在两天内进行核连缀程序。在大多数情况下，首先使用斑点印迹或狭缝印迹设备将未标记的 DNA 固定在膜上，这种设备可以从很多公司购买到。然后通过细胞膜的裂解从目的细胞制备细胞核。在放射性标记 NTP 存在的情况下孵育细胞核，然后纯化放射性标记的转录本并与含有未标记 DNA 的膜杂交。经过过夜杂交，通常首先用 RNA 酶处理膜，以消化未退火的 RNA，之后洗涤膜，最后通过放射自显影或磷屏成像仪分析膜。对弱表达基因的分析可能需要 2～3 周的长时间曝光（放射自显影)。

第 1 和第 2 天：核连缀分析。

第 3 天：数据分析。

提纲

步骤 1：准备缓冲液和溶液（1 h)。

步骤 2：将未标记的 DNA 固定到膜上（3 h)。

步骤 3：收集细胞并准备细胞核（1 h)。

步骤 4：放射性标记新生转录本（4 h）并开始杂交（过夜)。

步骤 5：洗膜（4 h)。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

[α-^{32}P] UTP〈!〉

乙酸铵（2.5 mol/L，pH 7)

乙酸铵溶液（5 mol/L 乙酸铵)

氯化钙溶液（$CaCl_2$，5 mmol/L 和 10 mmol/L)〈!〉

目的细胞（10^7～10^8)

氯仿〈!〉

无 RNase 的 DNase

乙醇〈!〉

HEPES（1 mol/L，游离酸)

异丙醇〈!〉

氯化镁（$MgCl_2$，10 mmol/L)〈!〉

$MgCl_2$/$CaCl_2$缓冲液

10 mmol/L $MgCl_2$〈!〉

5 mmol/L $CaCl_2$〈!〉

Nonidet P-40（NP-40）裂解缓冲液
10 mmol/L Tris-HCl（pH 7.4）
10 mmol/L NaCl
3 mmol/L $MgCl_2$〈!〉
0.5% NP-40〈!〉

核冷冻缓冲液
50 mmol/L Tris-HCl（pH 8.3）
40%（V/V）甘油
5 mmol/L $MgCl_2$〈!〉
0.1 mmol/L EDTA

苯酚：氯仿〈!〉

磷酸缓冲盐（PBS）

质粒 DNA，双链（10 μg），或单链 DNA

蛋白酶 K 缓冲液
10% SDS〈!〉
50 mmol/L EDTA
10 mmol/L Tris-HCl（pH7.4）
3 mg/ml 蛋白酶 K（新鲜加入）〈!〉

限制性内切核酸酶（用于双链 DNA）

RNase A（可选）

5×连缀缓冲液
25 mmol/L Tris-HCl（pH 8.0）
12.5 mmol/L $MgCl_2$〈!〉
750 mmol/L KCl〈!〉
1.25 mmol/L ATP
1.25 mmol/L GTP
1.25 mmol/L CTP

乙酸钠（0.3 mol/L）

氢氧化钠（NaOH）溶液
1 mol/L NaOH〈!〉
3 mol/L NaOH〈!〉

6×及 2×SSC
20×贮液：
3 mol/L NaCl
0.3 mol/L 柠檬酸钠×$2H_2O$〈!〉
用 1 mol/L HCl 调节 pH 至 7〈!〉

TE（pH 8）
10 mmol/L Tris-HCl（pH 8）

1 mmol/L EDTA

TES/NaCl 溶液

10 mmol/L TES（pH 7.4）

10 mmol/L EDTA

0.2% SDS〈!〉

0.6 mol/L NaCl

设备

临床离心机

锥形管

斑点印迹设备

滤纸

盖革计数器

离心机

硝酸纤维素膜（0.45 μm）

相衬显微镜

磷屏成像仪

塑料袋

保鲜膜

闪烁计数器

狭缝印迹设备

真空炉

程序

步骤 1：准备缓冲液和溶液

步骤 2：将未标记的 DNA 固定于膜上

1. 对于狭缝印迹的每个狭缝或斑点印迹的每个斑点，使用约 10 μg 双链质粒 DNA 或单链 DNA（每个阳性和阴性对照样品用相似数量的 DNA）。

如果使用双链 DNA，首先用限制性内切核酸酶切割使其线性化。用苯酚：氯仿提取，并在 0.3 mol/L 乙酸钠存在的情况下用乙醇沉淀。

2. 将用于每个狭缝或斑点的 DNA 样本稀释到 400 μl 的 TE 缓冲液中。

2. 加入 40 μl 3 mol/L NaOH 并在室温下变性和孵育 30 min。

3. 加入 300μl 2.5 mol/L 乙酸铵（pH7）进行中和。

4. 在狭缝或斑点印迹设备上将样品加到 0.45 μm 硝酸纤维素膜上。根据制造商的说明抽真空。每个狭缝用 6×SSC（1 ml）漂洗。

使用杂交膜而不是硝酸纤维素膜可能更容易取得成功的结果，但是注意某些对 Northern 或 Southern 印迹有用的膜在用于核连缀实验时，可能会产生非常高的背景。

5. 杂交膜经空气干燥，并在真空炉中 80℃烘烤 2 h。

6. 用 2×SSC 洗涤硝酸纤维素膜。

步骤 3：收集细胞并准备细胞核

每个样品从 10^7～10^8 个细胞开始。

1. 对于培养的细胞，使用一次性锥形管，在临床离心机中，1500 r/min 离心 10 min沉淀细胞。

2. 弃去上清液。通过悬浮细胞并以 1500 r/min 离心 10 min，用冷的（4℃）PBS 清洗细胞沉淀两次。

细胞冷却会导致聚合酶暂停在活跃转录的基因上。

3. 用 5 ml 冷的 NP-40 裂解缓冲液重悬细胞沉淀，在冰上孵育 5 min 裂解细胞膜。

取几微升样品在相差显微镜下分析，以确证大多数细胞已经裂解。

4. 在临床离心机中，4℃下 1000 r/min 离心 10 min 沉淀细胞核，弃去上清液。

5. 用 500 μl 核冷冻缓冲液重悬细胞核，并存于－70℃备用。细胞核可在－70℃稳定存放几年。

步骤 4：放射性标记新生转录本并开始杂交

1. 在冰上解冻冷冻的细胞核，并轻拍试管混合内容物。

2. 将 225 μl 含细胞核的混合物加到 60 μl 5×连缀缓冲液中。

3. 每管加入 15 μl（150 μCi）浓度为 3000 Ci/mmol（10 mCi/ml）的［α-^{32}P］UTP，37℃孵育 15 min。

在这一步中，延伸转录本被用放射性标记，起始一般不会发生。

理想的情况下，应设置多个时间点（如 37℃下 0、10 min、20 min、30 min），以确证在实验所用时间范围内，放射性的掺入大约呈线性。然而，考虑到每个样品需要很大的放射量，这并不总是可取的，并且这可能只对某些特定基因来说是必要的，即在所分析的细胞样品中转录速率被认为变化仅几倍或更小的基因。

4. 加入 20 μl 10 mmol/L $CaCl_2$ 及 10 μl 1 mg/ml 的无 RNase 的 DNase，30℃孵育 5 min 以消化基因组 DNA 并减低其黏度。

5. 加入 35 μl 蛋白酶 K 缓冲液。

6. 37℃孵育 45 min 以消化细胞蛋白。

7. 用 400 μl 酚∶氯仿抽提 2 次，并用 400 μl 氯仿抽提 1 次。特别小心不要转移出处于沿水相界面的沉淀。通过离心抽提后，移除水溶液中的所有沉淀。

8. 用浓度为 5 mol/L 乙酸铵（约 350 μl）1∶1 稀释水相。向样品中加入等体积的异丙醇，－20℃沉淀 30 min。

9. 在离心机中以 14 000 r/min 离心 10 min。将上清弃入放射性废物中，上清中应该含有大多数的未掺入的放射性核苷酸。

10. 用 100 μl TE（pH 8.0）悬浮沉淀。

此时，该溶液含有新生的放射性标记转录本。在乙酸铵存在的情况下通过乙醇沉

淀，大多数未掺入的放射性核苷酸应该已经去除。因此，放射性标记 RNA 理论上已准备就绪，只待与硝酸纤维素滤膜杂交。对于丰富的转录本，可以直接进行杂交步骤。然而，在大多数情况下，进一步纯化标记的 RNA 可能更有利于减少滤膜上的背景放射性，这些放射性似乎是由细胞中的 DNA 和蛋白质捕获，并在前面的步骤中未被去除。下面的步骤用来进一步纯化 RNA 和减少背景。用氢氧化钠处理产生的 RNA，使转录本裂解成更小的片段，这也能降低背景。对于纯化 RNA 的一个替代程序，该程序使用更稀的 DNase 和蛋白酶 K 消化，随后进行 RNA 的 TCA 沉淀，见 Ausubel 等（1994，单元 4.10）。

11. 加入 100 μl 含有 10 mmol/L $MgCl_2$ 和 5 mmol/L $CaCl_2$ 的溶液。

12. 加入 10 μl 1 mg/ml 无 RNase 的 DNase，37℃孵育 5 min。

13. 冰上冷却 5 min。

14. 加入 50 μl 1 mol/L NaOH 使转录本裂解成更小的片段，冰上孵育 2 min。

15. 立即加入 77 μl 1 mol/L HEPES（游离酸）进行中和，再加入 340 μl 5 mol/L 乙酸铵和 700 μl 的异丙醇。

16. −20℃冷冻 1 h。在离心机中以 14 000 r/min 转速离心 10 min，沉淀 RNA。

17. 在实验台上室温干燥沉淀，然后用 100 μl 的 TE 重悬沉淀。

18. 滴 1 μl 溶液在滤纸上，并在闪烁计数器中计数（切伦科夫计数），此步应能获得 3×10^5 cpm/μl 的典型掺入。

19. 用 2×SSC 洗涤硝酸纤维素滤膜。

20. 将一份放射性标记的 RNA 在塑料袋或闪烁瓶里与适当的硝酸纤维素滤膜条进行杂交（用于杂交的每个硝酸纤维素滤膜条应含有目的 cDNA 印迹狭缝或斑点，以及适当的阳性和阴性对照 DNA）。对于每个杂交，尝试大约 6×10^6 cmp 和 1 ml TES/NaCl 杂交液，65℃杂交至少 24 h。

此处建议的杂交缓冲液和条件均来自 Ausubel 等（1994）。然而，杂交缓冲液和条件可能需要优化。一般情况下，有助于 Northern 印迹的任何杂交条件都可能有助于核连缀杂交，包括用来增强杂交的甲酰胺缓冲液和硫酸葡聚糖。

步骤 5：洗膜

1. 用 2×SSC 在 65℃下洗涤硝酸纤维素膜 1 h，重复一次。用盖革计数器监测膜。阴性对照 DNA 的杂交应检测不到放射性。

增加洗涤对减少背景来说可能是必需的，可以用含 0.1% SDS 的 2×SSC 进行洗涤。通过在其中一次 2×SSC 洗涤（37℃，30 min）中包含 RNase A（约 10 μg）也可以减少背景，之后再用不含 RNase 的 2×SSC 洗涤 1 次（去除 RNase）。RNase 会消化掉放射性标记转录本中与膜上未标记的 cDNA 未退火的区域。除了降低背景，如果该实验是在进行与目的基因特定区域相关的新生转录本的定量分析，RNase 处理可能很重要，因为转录本中只有与膜上 DNA 直接退火的部分可抵御 RNase 的消化。

2. 用保鲜膜包裹杂交膜并在胶片或磷屏成像仪屏幕上曝光。

若凝胶图像显示背景过高，则可进一步增加膜的洗涤次数来降低背景。

附加说明

1. 基于以下考虑，在分析实验成功之前，可能需要多次尝试。核连缀分析对于来自高度活跃基因的转录本的分析是相当简单的。然而，即使是高度活跃的基因，大量的放射性也是必要的，而且加入放射性之后还涉及许多操作。对于不太活跃的基因，该分析可能很困难，其中最困难的步骤是产生一个令人信服的信号-背景比。弱活性基因导致弱信号或不可检测的信号，并且该分析很容易受到高背景的影响。

2. 已经发现步骤 3 中所述的用于制备细胞核的 NP-40 裂解法在几种不同类型的培养细胞中产生了成功的结果。然而，制备细胞核的替代方法也已有描述并有可能产生更好的结果（即新生转录本中掺入更多的放射性标记核苷酸）。Dounce 匀浆和蔗糖梯度离心分离用于核连缀分析的细胞核的方法见 Ausubel 等（1994，单元 4.10）。使用洋地黄皂苷改变细胞膜通透性的方法见 Ucker 和 Yamamoto（1984）。

3. 当把未标记的双链 DNA，而不是未标记的单链 DNA 固定到硝酸纤维素膜上时，重要的是要知道，由此产生的放射性信号可能归因于反义转录。例如，考虑使用固定化双链基因片段分析可诱导基因。如果只在诱导样品中检测到信号，那么可以断定该基因是在转录起始水平被诱导（至少部分被诱导）。然而，如果在未诱导和诱导的样品中获得相似强度的信号，那么该结果必须谨慎解释。或许，正如该结果所暗示的，该基因不是在转录起始水平被调控。另外，该基因座上的反义转录发生在未诱导和诱导两种细胞中，从而干扰了正义转录的检测，而正义转录可能被诱导。为了获得更令人信服的结果，硝酸纤维素膜应固定单链 DNA。

4. 通过噬菌体 T7、T3 或 SP6 RNA 聚合酶产生的未标记的 RNA，可代替未标记的 DNA 固定在硝酸纤维素滤膜上，但背景敏感性问题大大增加，因此不推荐这个策略。

疑难解答

放射性核苷酸很少掺入新生转录本

可能原因：核制备方法不当。

解决办法：裂解后通过显微镜确认细胞膜的裂解。尝试制备细胞核或者改变细胞膜通透性的其他方法（见上述“附加说明”，步骤 2）。

可能原因：在纯化程序中，RNA 或许已经丢失。

解决办法：重复。在用 DNase I 和蛋白酶 K 初步消化之后，测验三氯乙酸（TCA）沉淀材料中标记的掺入。

在洗涤后的滤膜上保留高放射性背景

可能原因：杂交或洗涤的严紧性太低。

解决办法：尝试不同的杂交和洗涤条件，遵照 Northern 印迹杂交和洗涤的基本参数（Sambrook and Russell 2001，7.42；Ausubel et al. 1994，单元 4.9）；在清洗期间

增加 RNase A 的消化程度。

可能原因：放射性标记 RNA 纯化不充分，使得残留的黏性 DNA 或蛋白质将放射性标记 RNA 捕获到滤膜上。

解决办法：在纯化放射性标记 RNA 时，使用新鲜批次的 DNase I 和蛋白酶 K；增加 DNase I 和蛋白酶 K 消化的次数；尝试不同的纯化放射性标记 RNA 的方法［例如，文献 Ausubel 等（1994），单元 4.10 中所描述的 TCA 沉淀，1994，单元 4.10］。

在硝酸纤维素膜上获得弱信号或检测不到信号

可能原因：低效率的杂交或稀有转录本。

解决办法：尝试不同的杂交和洗涤条件，遵照 Northern 印迹杂交和洗涤的基本参数（Sambrook and Russell 2001，7.42；Ausubel et al. 1994，单元 4.9）；增加膜上未标记的 DNA 的量；增加加入杂交中的放射性标记 RNA 的量；减少杂交的总体积；延长杂交时间；降低洗涤的严紧性；在标记新生转录本时，提高放射性标记的核苷酸的浓度。

参考文献

Ausubel F. M., Brent R. E., Kingston E., Moore D. D., Seidman J. G., Smith J. A., and Struhi K. 1994. *Current protocols in molecular biology*. John Wiley and Sons, New York.

Banerji J., Olson L., and Schaffner W. 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell*33: 729-740.

Blais A. and Dynlacht B. D. 2005. Constructing transcriptional regulatory networks. *Genes Dev*. 19: 1499-1511.

Calame K. and Ghosh S. 1995. Factors regulating immunoglobulin transcription. In *Immunoglobulin genes* (ed. T. Honjo, F. W. Alt, and T. H. Rabbits), pp. 397-422. Academic Press, London.

Cheadle C., Fan J., Cho-Chung Y. S., Werner T., Ray J., Do L., Gorospe M., and Becker K. G. 2005. Control of gene expression during T cell activation: Alternate regulation of mRNA transcription and mRNA stability. *BMC Genomics* 6: 75.

de Laat W. and Grosveld F. 2003. Spatial organization of gene expression: The active chromatin hub. *Chromosome Res*. 11: 447-459.

Derman E., Krauter K., Walling L., Weinberger C., Ray M., and Darnell J. E., Jr. 1981. Transcriptional control in the production of liver-specific mRNAs. *Cell* 23: 731-739.

Elnitski L., Jin V. X., Farnham P. J., and Jones S. J. 2006. Locating mammalian transcription factor binding sites: A survey of computational and experimental techniques. *Genome Res*. 16: 1455-1464.

Ernst P. and Smale S. T. 1995. Combinatorial regulation of transcription II: The immunoglobulin μ heavy chain gene. *Immunity* 2: 427-438.

Ghosh S., Gifford A. M., Riviere L. R., Tempst P., Nolan G. P., and Baltimore D. 1990. Cloning of the p50 DNA binding subunit of NF-κB: Homology to rel and *dorsal*. *Cell* 62: 1019-1029.

Gillies S. D., Morrison S. L., Oi V. T., and Tonegawa S. 1983. A tissue- specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* 33: 717-728.

Greenberg M. E. and Ziff E. B. 1984. Stimulation of 3T3 cells induces transcription of the *c-fos* proto-oncogene. *Nature* 311: 433-438.

Kim J., Chu J., Shen X., Wang J., and Orkin S. H. 2008. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* 132: 1049-1061.

Linial M., Gunderson N., and Groudine M. 1985. Enhanced transcription of *c-myc* in bursal lymphoma cells requires continuous protein synthesis. *Science* 230: 1126-1132.

Lipson K. E. and Baserga R. 1989. Transcriptional activity of the human thymidine kinase gene determined by a method using the polymerase chain reaction and an intron-specific probe. *Proc. Natl. Acad. Sci.* 86: 9774-9777.

Lo K., Landau N. R., and Smale S. T. 1991. LyF-1, a transcriptional regulator that interacts with a novel class of promoters for lymphocyte-specific genes. *Mol. Cell. Biol.* 11: 5229-5243.

Lonard D. M. and O' Malley B. W. 2006. The expanding cosmos of nuclear receptor coactivators. *Cell* 125: 411-414.

Ohtsu M., Kawate M., Fukuoka M., Gunji W., Hanaoka F., Utsugi T., Onoda F., and Murakami Y. 2008. Novel DNA microarray system for analysis of nascent mRNAs. *DNA Res.* 15: 241-251.

O' Malley B. W., Qin J., and Lanz R. B. 2008. Cracking the coregulator codes. *Curr. Opin. Cell Biol.* 20: 310-315.

Powell D. J., Friedman J. M., Oulette A. J., Krauter K. S., and Darnell J. E., Jr. 1984. Transcriptional and post-transcriptional control of specific messenger RNAs in adult and embryonic liver. *J. Mol. Biol.* 179: 21-35.

Reddy S. T., Gilbert R. S., Xie W., Luner S., and Herschman H. R. 1994. TGF-β_1 inhibits both endotoxin-induced prostaglandin synthesis and expression of the TIS10/prostaglandin synthase 2 gene in murine macrophages. *J. Leukoc. Biol.* 55: 192-200.

Sambrook J. and Russell D. 1991. *Molecular cloning: A laboratory manual*, 3rd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sandoval J., Rodriiguez J. L., Tur G., Serviddio G., Pereda J., Boukaba A., Sastre J., Torres L., Franco L., and López-Rodas G. 2004. RNAPol-ChIP: A novel application of chromatin immunoprecipitation to the analysis of real-time gene transcription. *Nucleic Acids Res.* 32: e88.

Siebenlist U., Durand D. B., Bressler P., Holbrook N. J., Norris C. A., Kamoun M., Kant J. A., and Crabtree G. R. 1986. Promoter region of interleukin-2 gene undergoes chromatin structure changes and confers inducibility on chloramphenicol acetyltransferase gene during activation of T cells. *Mol. Cell. Biol.* 6: 3042-3049.

Su R. C., Brown K. E., Saaber S., Fisher A. G., Merkenschlager M., and Smale S. T. 2004. Dynamic assembly of silent chromatin during thymocyte maturation. *Nat. Genet.* 36: 502-506.

Treisman R. 1985. Transient accumulation of *c-fos* RNA following serum stimulation requires a conserved 5′ element and *c-fos* 3′ sequences. *Cell* 42: 889-902.

Ucker D. S. and Yamamoto K. R. 1984. Early events in the stimulation of mammary tumor virus RNA synthesis by glucocortocoids. Novel assays of transcription rates. *J. Biol. Chem.* 259: 7416-7420.

网络资源

http://www.gene-regulation.com/pub/databases.html; http://www.biobase-international.com/pages/index.php? id =transfac：TRANSFAC包含关于转录因子、经实验验证的转录因子结合位点，以及转录因子调控的基因的数据。TRANSFAC广泛汇编结合位点，可用于派生位置权重矩阵。该数据库须订阅使用。

http://jaspar.genereg.net：JASPAR核心数据库包含一套经过专门人员审核过滤的、非冗余的图谱，这套图谱由123个图谱组成，这些图谱源于对已发表的、经实验界定的多细胞真核生物转录因子结合位点的收集。JASPAR是一个开放式访问数据库。

3　转录起始位点的定位

要点

- 在开始启动子分析前确定转录起始位点的位置非常重要。
- 尽管已经通过计算或全基因组实验方法预测了大量基因的转录起始位点，但个别实验验证通常是必要的。
- Cap 依赖性 5′ RACE 方法是目前起始位点定位的首选方法。

引言和概述

分析在转录起始水平上调控的基因的基本出发点是鉴定其启动子，并表征负责启动子活性的顺式作用序列元件和反式作用蛋白。正如第一章中所描述的，启动子包括在起始位点（起点）附近的 DNA 序列元件，这些元件指导转录的活化或抑制。根据这个定义，就可以简单地通过定位转录起始位点鉴定启动子区的位置。

精确定位起始位点对随后启动子分析的成功非常重要。在不止一例实验中，研究者对假定的启动子进行了相当详细的研究，但仅发现研究的区域距离起始位点很远，并且对基因调控也没有作用。在某些情况下，该失误是由于在分析开始时没有确定起始位点。在其他情况下，是由于将起始位点定位到了错误的位置。

通过鉴定编码信使 RNA（mRNA）的 5′端可确定一个基因的转录起始位点。mRNA 5′端序列相当于转录起始处的 DNA 序列。鉴定真核生物 mRNA 5′端的最早的方法

建立于 20 世纪 70 年代，但是这些方法烦琐且费时，同时，这些技术也不太灵敏，只能用于研究丰度最高的细胞和病毒转录本。例如，为了鉴定兔子 α-珠蛋白和 β-珠蛋白的 mRNA 5′端，Lockard 和 Rajbhandary（1976）首先利用寡聚 dT 层析和十二烷基硫酸钠（SDS）蔗糖密度梯度离心从网织红细胞多核糖体中纯化出了特定的 mRNA。之后，通过高碘酸盐的氧化作用除去 5′ 帽子，以便用［γ-^{32}P］ATP 和 T4 多核苷酸激酶放射性标记 5′端。通过凝胶电泳纯化末端标记的 mRNA，并用核酸酶 T1 消化。之后，通过对 T1 消化产生的放射性标记寡核苷酸的酶分析，确定 5′端的核苷酸序列。

在上述技术及其他几个比较繁琐的技术出现后不久，又发展了两种简单而灵敏的 RNA 分析方法。这两种方法都是间接的，即检测放射性标记的 cDNA，而不是检测 RNA 转录本和核苷酸本身。第一种方法称为 S1 核酸酶法，是由 Arnold Berk 和 Philip Sharp（1977）创立的。第二种方法，现在称为引物延伸方法（见方案 3.1），是由 P. K. Ghosh 和 S. M. Weissman（Ghosh et al. 1978）首次提出的。后来，在 20 世纪 80 年代产生了另外两种间接技术，即核酸酶（RNase）保护（Zinn et al. 1983；Melton et al. 1984；见方案 3.2）和 5′ RACE（cDNA 末端的快速扩增，Frohman et al. 1988，Loh et al. 1989）。下文将介绍 RACE、引物延伸法和 RNase 保护及其主要优缺点。对 S1 核酸酶法只做简单的介绍，该方法现在很少使用，这主要归因于其他方法的优点和改进。

虽然上面介绍的这些方法提供了关于大量基因的转录起始位点的有价值的信息，但每一个方法都高度易受实验假象的影响，这是由一些特殊的技术缺陷造成的，如 mRNA 降解和反转录酶对转录 5′ 端不完整的延伸（见下文）。20 世纪 90 年代，在消除这些主要缺陷方面取得了巨大进展，这主要是通过对 5′ RACE 方法改进实现的。第一个改进的 RACE 策略包括 RNA 连接酶介导的 RACE［RLM-RACE，也称为寡聚加帽（oligo-capping），Maruyama and Sugano 1994；Schaefer 1995］和 SMART（5′ 端 RNA 转录本开关机制）RACE（原本被称为 CapFinder；CLONTECHniques 1996）方法（见专题 3.1）。RLM-RACE 方法仅用于检测在 5′ 端有 7-甲基鸟苷酸帽子结构的转录本。由于仅在真正含有 RNA 聚合酶 II 转录本的 5′ 端才有这样的帽子结构，这种方法不会因为 RNA 的降解或反转录暂停而产生假阳性的 5′ 端。SMART RACE 方法也能消除对反转录暂停造成的假阳性 5′ 端的检测，但仍易受由 RNA 降解产生的假阳性 5′ 端的影响。

在证明 RLM-RACE 和 SMART RACE 成功用于单个基因分析之后，又发展了其他方法，如帽子俘获（CAPture）（Edery et al. 1995）和 CAP 捕手（CAP trapper）（Carninci et al. 1996），用于以分离和测序真核基因组所有基因全长 cDNA 为目标的大规模 cDNA 项目。CAPture 方法涉及含 5′ 帽子 mRNA 的亲和纯化，其中 5′ 帽子 mRNA 的捕获是通过它们与固定化鼠帽子结合蛋白 eIF4E 变体的结合实现的（Edery et al. 1995）。CAP 捕手方法依赖于第一链 cDNA 合成后 mRNA 帽子结构中生物素的直接化学引入，紧接着通过使用链亲和素包被磁珠纯化生物素化的 mRNA/ cDNA 双链（Carninci et al. 1996）。然后，将接头连接到 cDNA 3′端，以充当第二链 cDNA 合成的模板。一个新近开发的方法，即基因表达的帽子分析（CAGE）（图 3.1），将 CAP 捕手

与基因表达的系列扩增（SAGE）技术相结合，用于转录起始位点的高通量研究（Shiraki et al. 2003；Kodzius et al. 2006）。通过“配对末端双标签”的产生和基于SAGE的测序，对这种方法学已经进行了进一步改进，使其可以同时鉴定cDNA的5′端和3′端［被称为“基因组鉴定签名”（GIS）］（Ng et al. 2005）。一个相关的技术，即基因组签名克隆（GSC），已经被开发出来，实现了稀有cDNA的优先分离（见http://genome.gsc.riken.jp/）。使用5′ SAGE也已经进行了高通量研究，该研究将SAGE与寡核苷酸捕获策略相结合，用于加帽mRNA的选择（Hashimoto et al. 2004；Kasai et al. 2005）。最新的进展是对通过寡核苷酸捕获制备的常规文库进行Solexa测序（Wakaguri et al. 2008）。

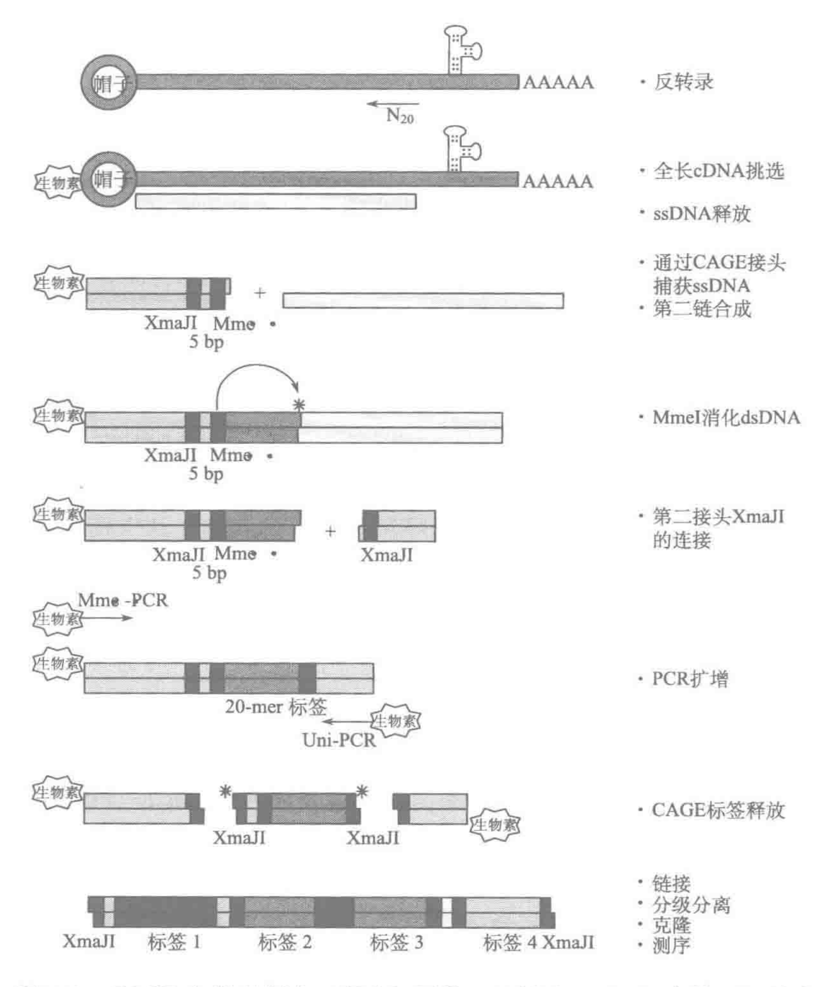


图3.1 CAGE文库的制备（经MacMillan Publishers Ltd. 允许，重绘自Kodzius et al. 2006）。

大规模制备和表征全长cDNA高度富集的cDNA文库，使人们获得了含有一个生物大量基因的候选转录起始位点的数据库。例如，通过对利用CAP捕手、CAGE、GIS

和 GSC 方法制备的 cDNA 文库的测序，RIKEN 基因组科学中心（RIKEN Genomic Sciences Center）（http://genome.gsc.riken.jp/）已经产生了一个详尽的鼠全长 cDNA 百科全书库。转录起始位点数据库（DataBase of Transcriptional Start Sites）（DBTSS；http://dbtss.hgc.jp/）包含一个人类基因起始位点的汇编，这些起始位点是通过对寡聚加帽方法（Wakaguri et al. 2008）制备的 cDNA 文库进行 Sanger 和 Solexa 测序获得的。另一个例子是伯克利果蝇基因组计划（Berkeley Drosophila Genome Project）EST（表达序列标签）集合。在这个集合中，大部分 EST 序列来自于帽子捕手方法制备的 cDNA 文库；EST 序列是利用一个和相邻载体序列杂交的引物对 cDNA 5′ 端测序获得的（Stapleton et al. 2002）。现在，大多数基因组数据库对基因组的注释尽可能多地使用 cDNA 序列信息，这些 cDNA 序列信息都来自于用大量富集全长 cDNA 的方法所制备的文库。

含有全长 cDNA 序列的数据库的获得提示了鉴定目的基因转录起始位点的适宜出发点，该出发点就是在相关生物的数据库中搜索该基因（基因组和 cDNA 数据库太多了，不能在本书中逐一列举，但很多数据库在核酸研究杂志的年度数据库期中会给予定期报告）。在这样的数据库中所描述的转录起始位点准确性相当高。但是，重要的是要获得另外的证据，表明从这样的数据库中提取的转录起始位点是有效的。这样的证据可能由起始位点的独立实验验证构成，或可能来自周围序列显而易见的属性，如下所述。获得另外的证据是必要的，因为尽管 CAP 捕手、CAGE 和寡聚加帽方法会产生大量的全长 cDNA，但它们并非万无一失。例如，利用链亲和素包被的磁珠的常用的 CAP 捕手和 CAGE 方法取得了效果，但从降解的非帽子化的 RNA 中分离全长的加帽 RNA 还不能尽善尽美（Sugahara et al. 2001）。此外，寡聚加帽方法需要用 RNA 连接酶将一个接头连接到 mRNA 的 5′端，而 RNA 连接酶显示序列特异性，并可能由此引入偏差或丢失某些重要的起始位点。

实验策略

初步考虑

如上所述，对于目的基因而言，有助于鉴定转录起始位点的恰当出发点，就是提取目的生物全长 cDNA 数据库中可能可用的信息。检查预测的起始位点周围的 DNA 序列可能会大大增加信心，相信预测的起始位点是有效的，或者可能怀疑注释的起始位点是不正确的。特别令人关注的是存在共有核心启动子元件、存在 CpG 岛、经染色质免疫沉淀（ChIP）评定的转录前起始复合体的结合，以及存在翻译起始密码子及其随后的处于预测转录起始位点合理距离处的预期开放可读框。

最重要的是，在预测的起始位点的上游 25～30 bp 存在一个共有 TATA 框，或者存在一个与预期位点重叠的共有 Inr 元件（见第一章，Smale and Kadonaga 2003），这些均可以大大增加利用数据库结果有效性的信心。一个启动子元件（DPE）下游的共有序列，也可以增加预测位点有效性的信心，虽然 DPE 的元件在哺乳动物基因中似乎不及在果蝇基因中常见（Smale and Kadonaga 2003）。值得注意的是，缺少 TATA、Inr

或 DPE 序列并不能排除起始位点在附近的可能性，因为许多基因不包含这些元件。相反，仅仅存在一个共有 TATA 或 Inr 元件也并不能令人信服地确立预测转录起始位点的有效性，原因是在基因组中只有部分 TATA 和 Inr 元件是有功能的，尤其是旁侧有适当的一群激活因子或抑制因子元件的那些 TATA 和 Inr 元件。

在预测转录起始位点附近 CpG 岛的存在（一个包含 CpG 岛的生物体）也能支持预测位点的有效性。现在已经知道，CpG 岛通常与大量哺乳动物启动子联系在一起（Gardiner-Garden and Frommer 1987；Slansky and Farnham 1996；Ioshikhes and Zhang 2000；Davuluri et al. 2001；Smale and Kadonaga 2003）。CpG 岛常与所谓的"持家"基因的启动子相关联，这样的基因一般很少表现出细胞类型特异性和发育调控。然而，许多细胞特异性的和发育调控的启动子也与 CpG 岛相关联。

然而，要注意的是现在还没有确凿地确定应当用于定义 CpG 岛的标准。早期的定义是一个 CpG 岛在长度上必须超过 200 bp，且必须拥有大于 50%的 GC 碱基对，以及观察到的 CpG 二核苷酸个数与预期的 CpG 二核苷酸个数的比值至少为 0.6（Gardiner-Garden Frommer 1987）。这个比值小于 1.0，这是因为 CpG 二核苷酸从哺乳动物的基因组已经耗尽，导致遍及大多数基因组的观察到的与预期的 CpG 二核苷酸比例非常低（约 0.1）。然而，其他研究表明，只有长度超过 500 bp、含有大于 55%GC 碱基对，并至少有 0.65 的观察与预期 CpG 二核苷酸比值时，这个区域才应被定义为 CpG 岛（Takai and Jones 2002）。造成这种更严谨的定义的实际原因之一是该定义使研究人员能够在全基因组研究中从 CpG 岛定义中排除 Alu 重复。然而，符合 Gardiner-Garden 和 Frommer（1987）提出的不太严谨的定义的区域，有可能拥有与更严谨定义的 CpG 岛的特征类似的特征。有各种基于网络的程序可用于确定一个 DNA 区域是否包含 CpG 岛，包括 CpGPlot（http://bioweb.pasteur.fr! seqanal/interfaces/cpgplot.html）、CpG Island Searcher（http://www.uscnorris. com/cpgislands2/cpg.aspx）、CpGcluster（http://bioinfo2.ugr.es/CpGcluster/）和 WebGene（http://125. itba.mi.cnr.it/cgi-bin/wwwcpg.pl）。

如果预测一个目的基因的转录起始位点在 CpG 岛内或紧邻 CpG 岛，那么这个起始位点或多个起始位点可能就在这附近。重要的是，要强调在一个 CpG 岛启动子中准确地鉴别精确的起始位点是困难的，原因包括以下两点。首先，CpG 岛启动子往往包含多个起始位点，分布在 100 个碱基对或 100 多个碱基对中。其次，高 G/C 含量增加了实验假象的可能性，这种实验假象与异常的 RNA 二级结构有关，它可以影响起始位点定位方法的成功。然而，因为准确定位转录起始位点的首要原因是要鉴定目的基因的启动子，所以了解启动子紧靠 CpG 岛会有助于实现这一目标（Bock et al. 2007）。

目的区域包含一个真实转录起始位点的其他证据，可从公布的全基因组启动子的定位研究中获取，这些研究使用 ChIP-芯片（ChIP-chip）或 ChIP-测序（ChIP-Seq）方法来确定与 RNA 聚合酶Ⅱ或转录前起始复合物的其他组件，如 TATA 结合蛋白（TBP）或 TBP 相关因子 1（TAF1）结合的 DNA 区域。分析老鼠和人类基因组的大量研究已经发表，这可能对这一目标的实现是有用的（如 Kim et al. 2005；Barrera et al. 2007；Zhao et al. 2007）。因为目的基因周围的全基因组 ChIP 结果的准确性无法得知，强烈

建议用独立的 ChIP 实验检查转录前起始复合物与跨越预测起始位点的基因组区域的结合情况。结果也许显示前起始复合物紧靠预测的起始位点形成，这为起始位点的有效性提供了额外的证据。

预测的起始位点也可能揭示存在一个下游翻译起始密码子（ATG），该密码子被一个 Kozak 一致序列（GCCACC ATGG；Kozak 1996）所包围，并且随后是一个开放可读框（ORF）。一个附近翻译起始密码子的存在可以增加预测转录起始位点有效性的信心。在脊椎动物基因中，从转录起始位点到 ATG 的平均距离一般为 100～150 个核苷酸（Pesole et al. 2002；Lynch et al. 2005）。此外，翻译通常在第一外显子中启动，尽管在许多基因中，翻译在其他外显子中启动。一项生物信息学研究显示 61%的人类基因的翻译是从第一外显子中开始的，其他 39%的基因是从完全非编码第一外显子中开始的（Davuluri et al. 2001）。这项研究还表明，含有 ATG 的第一外显子平均长度为 348 bp，而非编码第一外显子平均长度为 151 bp，此信息可以帮助研究者评估预测的转录起始位点是否有效。

最后，对预测的转录起始位点附近的基因组序列进行计算分析可能是有用的。已经发展了鉴定启动子、核心启动子、第一外显子和转录起始位点的大量计算方法（Zhang 1998，2007；Scherf et al. 2000；Davuluri et al. 2001；Bajic et al. 2006；Zhao et al. 2007），虽然这些方法不能完全代替对转录起始位点和启动子的实验验证，但它们依靠对基因组序列各种特征的检查正在迅速增加其效用。用于确定转录因子结合位点位置的数据库（如 TransFac），可能通过揭示附近序列与已知调控因子结合位点的同源性（见第 1 章），增加对预测起始位点有效性的信心。虽然已知转录因子的共有序列的存在能够增加预测起始位点有效性的信心，但这个信息必须再次谨慎对待，实际上，任何假定启动子序列的计算分析都会揭示已知转录因子的假定结合位点。正如上面对 TATA 和 Inr 元件的描述，功能相关的位点是那些相对于一群其他转录因子的结合位点恰当取向的位点，从而产生一个控制区，该控制区胜任核小体重塑和刺激含有 RNA 聚合酶 II 的前起始复合物的形成（见第 1 章和第 6 章）。

虽然在预测的转录起始位点附近的基因组 DNA 的特征可能导致相当大的信心，相信预测的位点是有效的，但许多研究者认为独立的实验验证是必要的。下面讨论用于鉴定和确证转录起始位点的 4 个实验方法。

快速扩增 cDNA 末端（RACE）

优缺点

强烈建议将 5′ RACE（专题 3.1）作为定位转录起始位点的首选方法，这主要是因为新的 RACE 策略的建立，这些新策略以比传统 5′ RACE 更高的成功率鉴定转录本真正的 5′端。这些新的 RACE 方法比原来的 RACE 方法或其他的起始位点定位方法更不易受实验假象的影响。例如，传统 5′ RACE（和引物延伸分析）的一个局限性是由于稳定 RNA 二级结构的存在，反转录酶常常不能延伸到 mRNA 的 5′端。在传统 RACE 和引物延伸程序中，这些不完整的 cDNA 都可能被错误地指定为潜在的转录起始位点。较新的 RACE 程序包括导致最终产物几乎全部来自全长 cDNA 的步骤。从理论上讲，

FirstChoice RLM-RAC 试剂盒（Ambion 公司）比 SMART RACE 试剂盒（Clontech 公司）拥有新的优势：RLM -RACE 试剂盒只从含 5′ 帽子结构的 mRNA 分子产生一个可检测产物。而 SMART RACE 试剂盒只有当反转录酶成功延伸到转录本的 5′端时，才能得到一个产物，但因为它不要求有 5′ 帽子，会产生来自降解 RNA 的产物。然而，RLM-RACE 试剂盒的一个潜在的缺点是，它需要使用 RNA 连接酶进行 RNA 接头的连接，而 RNA 连接酶具有显著的序列偏好（Schaefer 1995）。因此，对某些 mRNA 而言，RLM-RACE 试剂盒可能比其他的试剂盒效果要好。此外，通过 RNA 样品的初始磷酸酶处理，Ambion 公司的程序会抑制对降解 mRNA 的检测；如果这种磷酸酶反应是不彻底的，降解产物仍然会被检测出来。其他公司也销售 RACE 试剂盒（如 Invitrogen 公司 2.0 版本的 5′ RACE 试剂盒；罗氏应用科学部第二代 5′ /3′ RACE 试剂盒）。遗憾的是，据我们所知，没有人曾独立比较过以各种试剂盒和发表的方法取得的结果的质量。

RACE 试剂盒除了具有富集延伸到一个转录本 5′端的 cDNA 的独特能力外，另一个明显的优势是其具有很高的灵敏性，因为它是唯一依赖于 PCR 扩增的方法。对于那些不能有效转录的或其他方法已证明不能定位转录起始位点的基因，这种灵敏性提供了一个独特的机会。5′ RACE 最后一个主要优点是扩增产物可以被克隆和测序。下面介绍的其他三个方法，内含子和外显子-内含子联结点可导致误导性结果（见专题 3.5）。此外，如果第一个外显子在已知外显子上游远端且难以鉴定，其他三个程序将不能提供有关其位置的重要信息。相反，如果将 5′ RACE 产物克隆到载体并测序（或将 RACE 产物直接测序），该序列将有可能揭示第一外显子和转录起始位点的位置。

因为我们强烈建议研究人员使用商业试剂盒进行 5′ RACE 实验工作，所以不包括具体的方案和实验建议。相反，建议研究人员查阅制造商的产品说明书。

专题 3.1 帽子依赖性 RACE 程序

基本 RACE 程序（图 3.2）（Frohman et al. 1988；Frohman 1995）从引物延伸步骤开始，将未标记寡核苷酸引物与目的 mRNA 的前 100 或前 200 个核苷酸退火，利用反转录酶将之延伸到 RNA 5′ 端。然后，利用聚合酶链反应（PCR）扩增 cDNA 产物，cDNA 产物的 3′ 端与 mRNA 5′ 端相对应，并因此和转录起始位点相对应。由于 3′ 端的 cDNA 未确定，直到添加确定序列的核苷酸后，才能进行 PCR。在最初的 5′ RACE 程序中，这是通过用末端转移酶添加一个同源多聚体序列或通过连接一个确定序列的单链寡核苷酸来完成的。在确定的序列添加到第一链 cDNA 的 3′ 端后，与确定序列互补的引物，以及比用于最初 cDNA 合成步骤的引物稍微靠内的第二个引物，可用于 cDNA 扩增。扩增产物通常插入到一个载体，分离几个单个质粒克隆，并确定每个克隆的序列。通过对来自 PCR 混合物的许多克隆的测序，可以推断出假定 RNA 起始站点的位置。另一种方法是用 5′ 端标记的引物进行最终的 PCR 分析，然后在变性聚丙烯酰胺凝胶上确定产物大小。通过与适当的分子质量标准物比较产物的大小，就可以预测转录起始位点的位置。

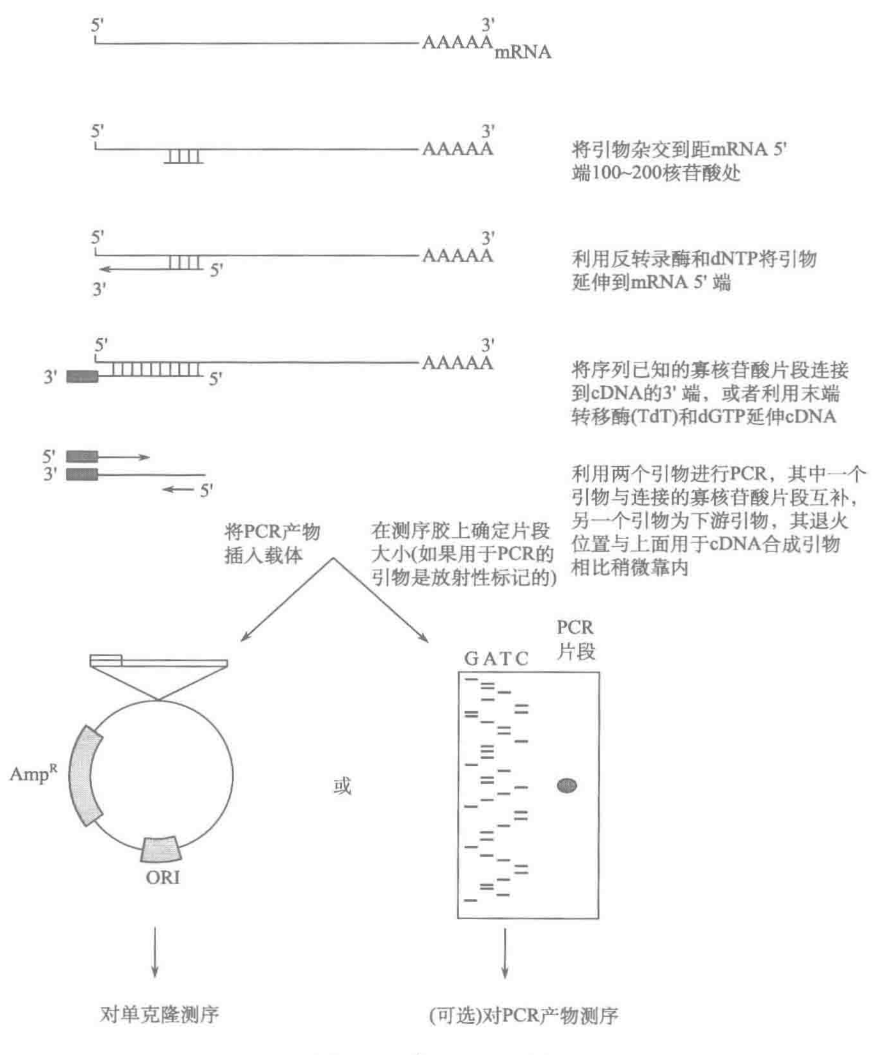


图 3.2　5′ RACE 程序。

强烈建议使用可富集全长 cDNA 或优先检测具有 5′ 帽子结构的 mRNA 的 5′ RACE 的改进版。在 FirstChoice RLM-RACE 试剂盒（Ambion）中，mRNA 样品首先用小牛肠碱性磷酸酶处理，从降解的 RNA 中去除 5′ 磷酸（全长戴帽 mRNA 具有碱性磷酸酶抗性）。之后，通过烟草酸焦磷酸酶的处理从全长 mRNA 中将 5′帽子去除，这种处理方法在 mRNA 的 5′ 端保留单磷酸。这样，现在就只有全长 mRNA 才含有 5′ 磷酸。使用 RNA 连接酶将一个 RNA 接头连接到 mRNA 上，并且将只连接到含有 5′ 磷酸的 RNA 分子上。因此，这种方法能大大富集全长戴帽 mRNA。在这一点上，一个基因特异性引物与目的 mRNA 杂交，利用反转录酶（延伸将通过 RNA 接头）延伸到 RNA 分子的 5′ 端。然后，可以使用第二个基因特异引物和含有与 RNA 接头相同序列的引物进行 cDNA 第一链的 PCR 扩增。

Clontech 公司的 SMART RACE cDNA 扩增试剂盒（SMART RACE cDNA Amplification Kit）采用了不同的策略以富集全长 cDNA。在此试剂盒中，第一链 cDNA 合成是按原始方案进行的。然而，当反转录酶达到 RNA 分子的 5′ 端时，其天然末端转移酶活性会添加数个额外的脱氧胞苷残基。Clontech 公司开发了一个帽子转换（CapSwitch）寡核苷酸，该寡核苷酸包含一段可与这段脱氧胞苷退火的脱氧鸟苷。使用帽子转换寡核苷酸作为模板可使反转录酶继续延伸，从而给每个第一链 cDNA 的 3′ 端增加了确定的序列。由于只有当反转录酶达到 RNA 分子的 5′ 端，并且在延伸期间不停顿时，才能添加一段脱氧胞苷，这个程序富集全长 cDNA。该程序的下一步是使用基因特异性下游引物和与帽子转换寡核苷酸区互补的第二个引物，进行 PCR 扩增。

引物延伸

优缺点

当尝试鉴定一个新基因的起始位点时，以前推荐使用的第一种方法是引物延伸（专题 3.2）。不过，随着 5′ RACE 程序的改进，情况有所改变。然而，相对于 RNase 保护和 S1 核酸酶分析，引物延伸的操作更为简单快速。引物延伸可在大约 5h 内完成，需要的特殊试剂仅仅是 mRNA 样品和与基因内一段序列互补的放射性标记的核苷酸引物。由于使用与 mRNA 有效杂交的引物，因此这种方法特别灵敏，但还是没有 5′ RACE 灵敏。引物延伸的另一个好处是，它为常规监控基因表达提供了一种可替代实时 RT（反转录）-PCR 和 Northern 印迹的富有吸引力和定量的方法。

引物延伸方案的一个缺点是难以找到对一个新基因有效的引物。而且，由反转录的过早中止（归因于 RNA 二级结构）造成的背景条带经常出现，使得明确确定起始位点的位置变得困难。对于有些 mRNA，如与 CpG 岛相关联的 mRNA，反转录效率可能特别低。由于这些局限性，我们推荐用引物延伸方法获得的起始位点定位结果应使用本章讲述的其他方法加以确证。

专题 3.2 引物延伸

引物延伸方案（图 3.3 和图 3.4）始于 5′ ^{32}P 标记的引物，引物通常是人工合成的寡核苷酸，与预期 5′端的下游 50～150 个核苷酸的 mRNA 序列互补。在根据经验确定的反应条件下，过量的放射性标记引物与从目的细胞中分离的总 RNA 或寡聚（dT）纯化的 mRNA 中的特异 RNA 分子退火。将反转录酶、脱氧核苷三磷酸和适宜的缓冲液组分加入引物-mRNA 杂交体中，催化引物延伸到 mRNA 5′ 端。所产生的放射性标记 cDNA 产物用变性聚丙烯酰胺凝胶电泳分析。通过与邻近的测序梯状条带或分子质量标准物的比较，在凝胶中检测的条带的大小提供了一个从合成寡核苷酸 5′端到 mRNA 转录起始位点的距离的测量。如果标记的 cDNA 产物在凝胶的分辨率范围内，则可以以 ±1 个核苷酸的精确性确定转录起始位点。

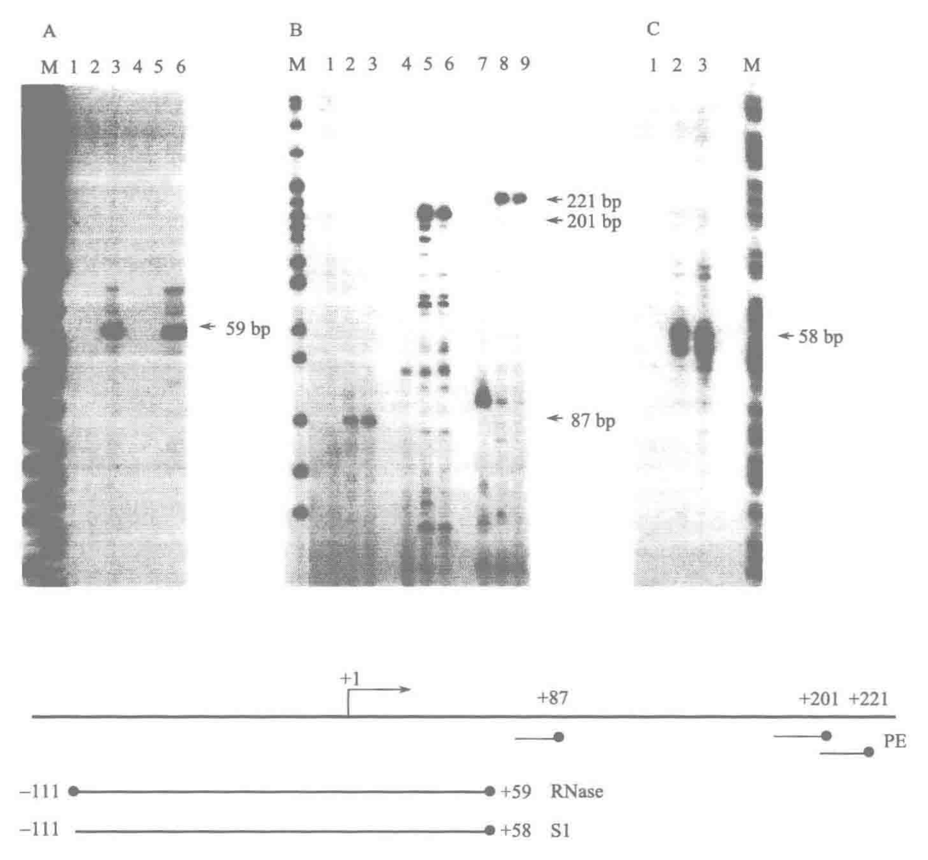


图 3.3　来自 RNase 保护、引物延伸、S1 核酸酶实验定位鼠末端转移酶（TdT）基因转录起始位点的结果。（A）以 TdT 基因表达和不表达的细胞系中的胞质 RNA 进行的 RNase 保护分析。来自两个表达 TdT 基因的细胞系的胞质 RNA（泳道 3 和 6）；来自 4 个不表达 TdT 基因的细胞系的胞质 RNA（泳道 1、2、4 和 5）。标记的探针从核苷酸＋59 延伸到-111。预期产物是 59 个核苷酸。图中也显示了测序分子质量标准物（M）（经允许转载自 Lo et al. 1991）。（B）以 TdT 基因表达和不表达的细胞系胞质 RNA 进行的引物延伸分析。来自两个表达 TdT 基因的细胞系的胞质 RNA（泳道 2、3、5、6、8 和 9）；来自一个不表达 TdT 基因的细胞系的胞质 RNA（泳道 1、4 和 7）。使用三个不同的引物，它们的 5′ 端分别与预期的起始位点下游 87bp（泳道 1～3）、201 bp（泳道 4～6）、221bp（泳道 7～9）处杂交。分子质量标准物（M）是来自质粒 pBR322 的 5′ 端标记的 MboI 限制性片段。（C）以 TdT 基因表达和不表达的细胞系中胞质 RNA 进行的 S1 核酸酶分析。来自两个表达 TdT 基因的细胞系的胞质 RNA（泳道 2 和 3）；来自一个不表达 TdT 基因的细胞系的胞质 RNA（泳道 1）。标记的探针从核苷酸＋58 延伸到-111，预期产物是 58 个核苷酸。测序分子质量标准物显示为 M（B，经 Elsevier 允许，转载自 Smale and Baltimore 1989）。

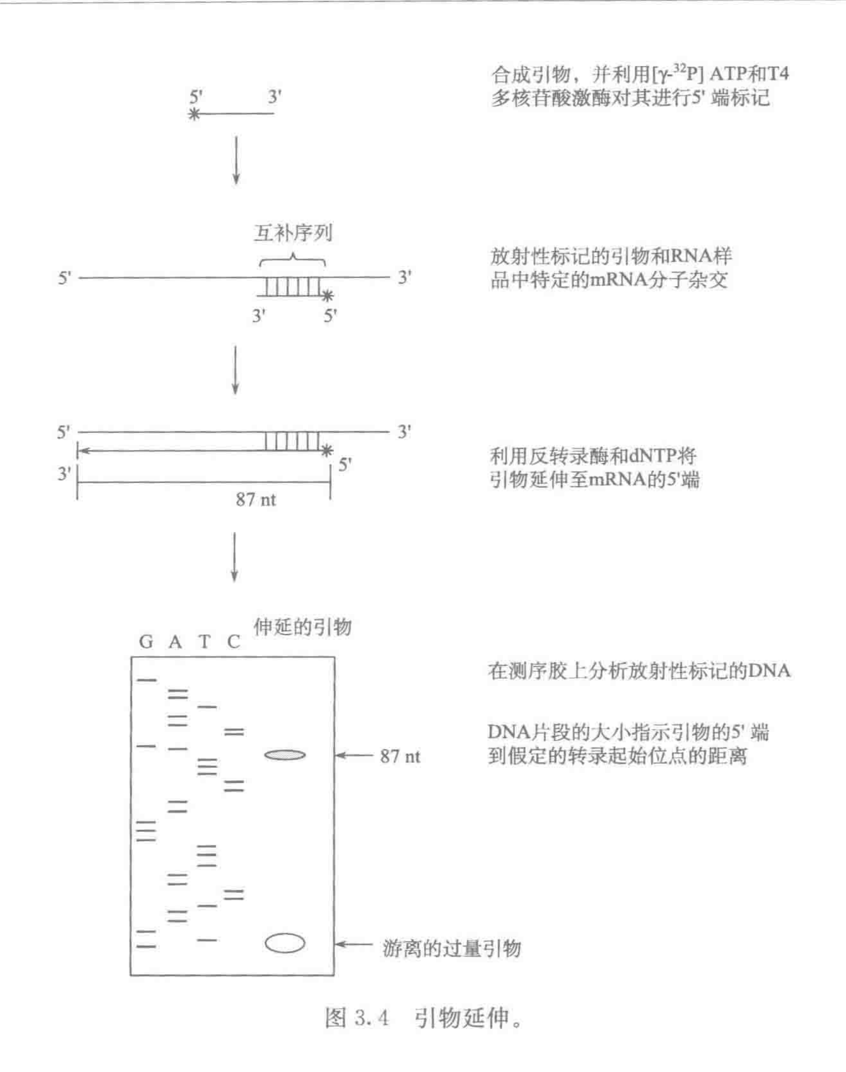


图 3.4　引物延伸。

寡核苷酸引物设计

许多不同长度的寡核苷酸引物都能成功地用于引物延伸。在人工合成的寡核苷酸广泛可用之前，来自 100 或 100 多个核苷酸的限制性片段的单链引物曾被成功地应用。但是现在已很少使用这种长度的引物，因为现今人工合成寡核苷酸既便宜又容易获得。9 个核苷酸的合成寡核苷酸就足以进行杂交和延伸，但在引物延伸法中并不使用这么短的寡核苷酸，因为这么短的寡核苷酸的杂交缺少特异性和高效性。20～25 bp 的寡核苷酸一般能产生很好的结果。有多个软件程序可用于设计研究者心想的具有一定长度和解链温度（T_m）的引物。“引物 3（Primer3）”程序（http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi）就是这样一个被许多研究者首选的程序。可商业性获得的程序包括来自 DNASTAR 公司的“引物选择（PrimerSelect）”（http://www.dnastar.com/products/primerselect.php），以及来自 Molecular Biology Insight 公司（http://

www. oligo. net/）的“寡聚引物分析软件（Oligo Primer Analysis Software）”。

尽管推荐使用计算机程序来设计引物，但利用不同引物所获得的实际成功率并不相同。如果利用一个最初的引物所获得的结果不很理想，应测验与 mRNA 5′ 端附近的其他序列互补的一个甚至数个另外的引物。目前尚不清楚为什么具有相似 T_m 值的不同引物所获得的结果差别很大，但引物标记效率的差异或 mRNA 的二级结构可能对此产生一定的作用。

为了精确定位转录起始位点，引物的 5′ 端最好在距 mRNA 起始位点 50～150 个核苷酸的位置退火。如果和相邻的 DNA 测序梯状条带一起电泳，这个长度可以使产生的 cDNA 产物在变性聚丙烯酰胺凝胶上被精确测量。小于 50 bp 的 cDNA 产物是不受欢迎的，因为反转录酶反应有时在延伸非常短的距离后终止或暂停，结果产生比引物长10～20 bp 的强背景条带。另外，反转录酶有时会合成引物的一个拷贝，导致所谓的“引物二聚体”人造物。如果合成的 cDNA 至少是 50 个核苷酸（如 20 个核苷酸的引物加上 30 个核苷酸的延伸），就会与大部分背景条带很好地分开。经常可获得长于 150 bp 的 cDNA 产物，但由于暂停和提前终止，反转录酶延伸的效率会随着延伸距离的增加而降低。而且，由于大片段在变性聚丙烯酰胺测序胶中的分辨率降低，因此延伸产物越长，越难确定其精确长度。

引物退火和反转录

对每个被试引物，必须根据经验确定退火条件，因为根据核苷酸含量预测的温度并不总是精确的，本章后面所列的引物延伸程序中，退火反应在 250 mmol/L KCl 中进行，在这一盐浓度下，建议测试的退火温度是 37℃、45℃、60℃和 68℃。还可以将退火反应在一温度块中加热到 68℃然后使其慢慢冷却到室温。

本章中的引物延伸反应采用在大肠杆菌中克隆和表达的莫洛尼氏鼠白血病毒（Moloney murine leukemia virus，MMLV）反转录酶。MMLV 反转录酶含有 RNase H 活性，可导致背景增强，但经改造的不具 RNase H 活性的反转录酶可从 Invitrogen 公司购买（如 SuperScript II RT 和 SuperScript III RT）。SuperScript III RT 含有另外的突变，这些突变会增加热稳定性和半衰期，结果使延伸反应可在高温下进行，从而减少与 RNA 二级结构相关的问题。当然也可利用其他反转录酶进行引物延伸反应，如鸟类反转录酶，但反应条件不同，对 pH 的变化更为敏感。

数据分析范例

对于图 3.3B 中所示的末端转移酶（TdT）引物延伸的例子，测试了三种不同的引物（Smale and Baltimore 1989）。一个引物的 5′ 端在转录起始位点下游 87 bp 处杂交，其他两个引物的 5′ 端在起始位点下游 201 bp 和 221 bp 处杂交。在凝胶中迁移非常迅速的大量过量的游离引物，已从该图片显示的放射自显影照片底部被除去了。为了确证该引物绝对过剩，在凝胶底部保留游离引物是有益的。以每个引物所观察到的最大延伸产物都定位于一个单一的转录起始位点。第二个和第三个引物的杂交位置比上述推荐的转录起始位点更远（201 bp 和 221 bp，而不是 50～150 bp）。然而，这些引物在本实验中被有效延伸至 mRNA 5′ 端。在延伸效率较低的其他实验中，全长的延伸产物丰度低，并伴随着较短的、部分延伸的产物（数据末出示）。利用三个不同引物获得的信号强度

尽管在凝胶上是相当的，但其实是有变化的，因为反应产物的量被归一化以产生在放射自显影上相似强度的条带。引物间的变化可能是由标记或退火效率的差异造成的。

对每个引物来说，研究人员测验了来自两种表达 TdT 基因的鼠细胞及一种不表达 TdT 基因的鼠细胞的总细胞 mRNA，其中后者作为阴性对照。使用来自相同物种的阴性对照意义重大，因为经常会观察到由于引物与无关 RNA 杂交造成的背景条带，这些条带在阴性对照中的存在表明它们与组织特异性目的转录本无关。用第二个和第三个引物时都观察到了两例这种类型的背景条带，其中一例在每种情况下都位于放射自显影的中间部位，而另一例位于放射自显影的底部。第三个引物在阴性对照泳道中产生了非常强的背景条带，而这一背景条带在阳性对照泳道中很少，在其他引物中未检测到。使用第二个引物时缺乏这种背景条带，表明这一背景条带是由不相关的转录本引起的，第三个引物偶然与该转录本杂交。很显然，这种转录本只在阴性对照细胞系中强烈表达。

其他明显的背景条带（主要以第二个引物）可能产生自反转录酶在全长 mRNA 延伸过程中的停顿。严重 mRNA 的降解，如上所述的与不相关转录本的杂交，或者多重转录起始位点的存在等，也能造成背景条带。引物延伸法的一个主要的局限性是背景条带十分普遍，就像以第二个引物所观察到的那些背景条带，因此很难确定哪条带对应于真正的转录起始位点，以及哪条带是实验假象。对于 TdT 基因而言，利用多个引物延伸，接着再用 RNase 保护及 S1 核酸酶分析确证，就可以对真正的 mRNA 起始位点作出精确定位。

RNase 保护法

优缺点

RNase 保护程序（专题 3. 3）被推荐作为引物延伸法最具吸引力的替代方法。由于引物延伸和 RNase 保护的原理不同，以及受不同类型实验假象的影响，转录起始位点的严格论证通常涉及两种技术的使用。如果通过两种方法鉴定了相同的位点，那么该位点有较高的可能性代表真正的转录起始位点。RNase 保护比引物延伸费时，但是，除了 5′ RACE 外，它通常是定位起始位点的最灵敏的方法。RNase 保护灵敏高的的原因主要有两个。其一，探针通常用高比活的 [α-^{32}P] UTP 合成，因此产生的探针中有 10%左右尿嘧啶残基被放射性标记，相比之下，引物延伸法的引物中每个分子平均不到一个放射性磷原子。其二，在 RNase 保护中形成的 RNA-RNA 杂交分子在热力学上是稳定的，比 S1 核酸酶分析中形成的 RNA-DNA 杂交分子稳定得多。这种稳定性能通过降低在核酸酶消化步骤中杂交分子的“呼吸”（瞬时解离）频率而增加灵敏性。

RNase 保护程序的主要缺点是很难为新基因建立该分析方法，因为杂交和消化条件必须依经验确定。而且，放射性标记探针容易受稳定二级结构形成的影响，稳定的二级结构可产生意外的 RNase 抗性产物。最后，高比活的放射性标记探针只能使用几天，因为它们趋向于经历快速辐射分解。相反，引物延伸分析中标记的寡核苷酸引物通常可使用近一个月。

探针制备

制备 RNase 保护探针比较容易。最简单的方法是使用基因组克隆或基因组 DNA 作

为模板，利用两个 PCR 引物（其中一个引物包含噬菌体启动子序列），通过 PCR 扩增跨越转录起始位点的基因组序列（见专题 3.3）。噬菌体启动子应该包括在下游引物中，以便基因组序列以反义方向结束。PCR 产物可以直接用来制备探针。

如果需要经常制备相同的探针，制备含噬菌体启动子和反义基因组 DNA 片段的质粒可能非常有利。为了实现这一策略，可以利用含噬菌体启动子的引物对基因组片段进行 PCR 扩增，然后将之插入到质粒载体中。或者，也可以将基因组 DNA 片段呈反义方向克隆到已含有噬菌体启动子的载体中。质粒的构建并无多大难度，但如果在插入载体前用 PCR 扩增基因组 DNA，必须对扩增产物进行测序，以确证没有引入突变；突变会破坏 RNA-RNA 双链的完整性，从而引起不受欢迎的 RNase 切割。正如所有的程序，必须进行转录起始位点可能位置相关的预测，以确定作为探针的基因组 DNA 片段的边界。如果证明预测是不正确的，将需要跨越不同区域的其他探针。

必须解决的一个关键问题是探针应该多长。一般来说，探针往转录区延伸得越长，分析的灵敏性就越高，原因是较长的探针分子掺入了更多的放射性磷原子。但是，较长的探针含有二级结构的可能性较大，这会导致背景信号和无效杂交。可以利用 Michael Zuker 的“mfold”算法计算不同区域出现二级结构的数量。通过 Zuker 实验室主页（http://frontend. bioinfo. rpi. edu/ zukerm/home. html）可访问此程序及其他二级结构预测程序。较长的探针分子可能还包括使噬菌体聚合酶不能有效延伸的 DNA 区域。因为获得了大量不完全转录本，通常有必要在较长的探针可使用之前对其进行凝胶纯化，但短探针分子（如小于 300 个核苷酸）在使用前不需要进行凝胶纯化。为了确定是否有必要进行纯化，通过在凝胶中电泳并自显影一小份放射性反应产物，可确定全长探针分子的大概比例。

专题 3.3　RNase 保护

RNase 保护程序（图 3.5）从掺入一种［α-^{32}P］NTP（通常是［α-^{32}P］UTP）制备均匀标记的 RNA 探针开始。RNA 探针用噬菌体 RNA 聚合酶（SP6、T7、T3）合成，该酶从插入大量普通质粒载体的特异性噬菌体启动子处开始转录。为了定位起始位点，质粒模板包含一个基因组 DNA 片段，该片段跨越被认为含有目的基因转录起始位点的区域。将这个基因组片段呈反义方向亚克隆到质粒中噬菌体启动子的下游，以便产生的 RNA 探针的 5′ 端的一部分序列将与目的 mRNA 互补。也可以采用 PCR 产生用于制备探针的 DNA 模板，其中一个 PCR 引物应含有噬菌体聚合酶启动子序列，然后扩增基因组序列。放射性标记的探针与从目的细胞中纯化的细胞质 mRNA 或总 mRNA 退火，杂交反应通常在甲酰胺缓冲液中进行数小时或过夜。然后将 RNase A 与 RNase T1 或其中之一加入到杂交反应中，这些核酸酶消化 RNA 分子上的单链突出部分，而 RNA-RNA 杂交分子对这些酶的切割具有抗性。这种抗性形成了该程序的概念基础，因为与特定 mRNA 分子退火的探针区域能抵抗消化。探针抗消化区域的长度对应于从探针 5′ 端到转录起始位点的距离。由于该探针是放射性标记的，可以在高分辨率的变性聚丙烯酰胺凝胶电泳上观察抗性片段的大小。

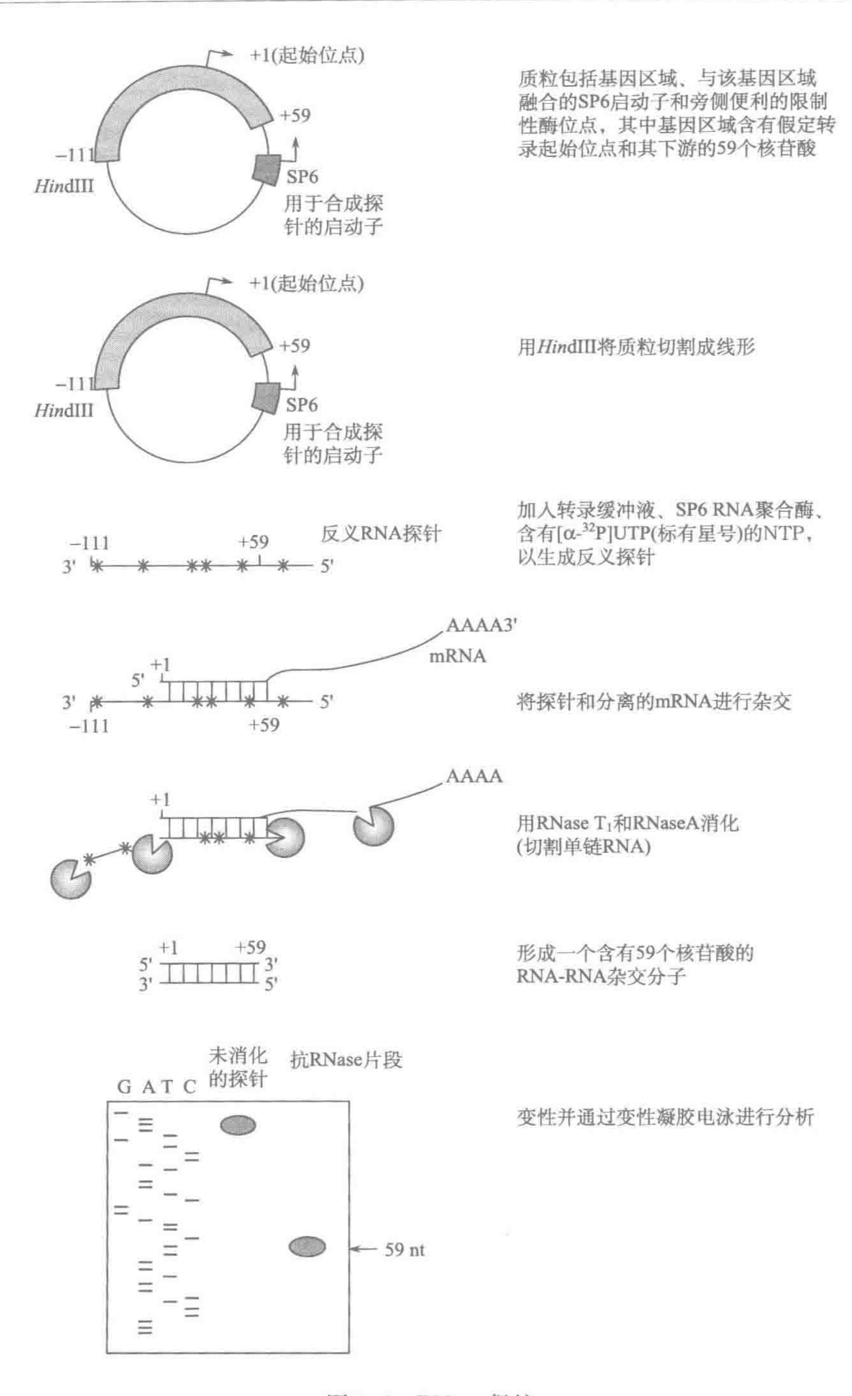


图 3.5 RNase 保护。

探针退火和 RNase 消化

上文已提到，退火和 RNase 消化条件必须依经验确定。不同探针的退火温度变化

很大，并且不仅仅取决于 RNA-RNA 双链的计算 T_m值。降解单链产物所需要的 RNA 酶的精确浓度也因探针不同而有很大变化。

为了简化对一个新基因发展 RNase 保护分析方法，可以利用商业试剂盒。Ambion 公司提供基本 RNase 保护试剂盒（RPA III）和为更快速杂交（HybSpeed RPA Kit）及使用粗制 RNA 制剂（Direct Protect Lysate RPA Kit）设计的改进试剂盒。这些试剂盒含有已经对其效价进行了滴定的 RNase 溶液，还包括制备对照探针的模板，对照探针对于确证试剂正确地发挥功能很有用。由于利用已精确测定效价的 RNase 贮液和对照探针，在建立 RNase 保护分析并解决 RNase 保护分析中的问题时，需要考虑的变量就更加有限。

对来自表达和不表达目的基因的相同物种的细胞系的 RNA 样品进行比较，也有助于分析方法的发展。用不表达基因的细胞 mRNA 所进行的反应中当然不会检测到相关的条带，从而使研究者很容易鉴定出一些不相关的条带。

数据分析范例

在图 3.3A 所示的例子中，从核苷酸－111 到＋59（Lo et al. 1991）的基因组序列制备分析 TdT 转录起始位点的 RNase 保护探针，并将该探针片段反向插入到 pSP72（Promega 公司）中 SP6 启动子下游。在探针制备之前，用限制性内切核酸酶在核苷酸-111 附近切割质粒，这样，在去除切割的模板之前，SP6 RNA 聚合酶将合成含 180 个核苷酸的放射性标记的产物。由于探针相对较短，不含 SP6 聚合酶的强终止或暂停位点，因此不需要进行凝胶纯化。放射性标记探针可与 6 个不同的鼠细胞系（其中只有 2 个细胞系表达 TdT 基因）的 RNA 样品杂交（泳道 3 和 6）。随后用 RNase A 和 RNase T1 酶切样品，消化杂交分子上的单链 RNA 凸出。当通过变性凝胶电泳分析时，经与分子质量标准物（Marker）比较确定，丰度最高的抗性产物迁移在大约 59 个核苷酸处。根据这个产物的大小，就可将 mRNA 5′端定位在引物延伸所确定的相同位置上。

在来自表达 TdT 的细胞系的两个泳道中，检测到了比主带略长或略短的较弱的条带。如果 RNase 保护是用于定位转录起始位点的唯一方法，那么将很难确定这些条带是否代表次要的转录起始位点。由于通过引物延伸在相同位置上没有检测到这些弱带，因此这些弱带可能是由其他原因引起的。最可能的是，较大的片段是由 RNase A 和 RNase T1 相关的事实引起的，这种事实是这两种酶都是内切核酸酶而不是外切核酸酶，它们并不总是在分开双链和单链 RNA 的最后一个磷酸二酯键上切割，而是在留下短凸出的位点上切割。比主带略短的条带最有可能产生于 RNA-RNA 双链末端的呼吸，同时 RNase 偶尔切割内部磷酸二酯键。这些结果表明，可以利用两种不同的分析方法来区分代表每种分析的真正 5′ 端及实验假象的条带。

S1 核酸酶分析

优缺点

与引物延伸或 RNase 保护程序相比，S1 核酸酶程序（专题 3.4）现在已很少使用。但是，如果采用其他方法分析一个特定基因比较麻烦时，该方法则是一个备选的方法。当使用 5′ 端标记的探针时，这对 S1 核酸酶分析非常常见，相对于 RNase 保护，其主

要优点是无背景条带，可能是因为异常切割的产物只有当它们包含放射性标记的 5′ 端核苷酸时才是可检测的。当使用 5′ 端标记的探针时，主要的缺点是该程序不太灵敏，因为探针分子具有平均不足一个放射性磷原子。另一个潜在的缺点是 RNA-DNA 双链分子对呼吸和随后单链区的 S1 核酸酶切割更敏感。如果杂交体含有长的富含 A-T 的区段，形成相对不稳定的碱基对，则呼吸会特别的麻烦。

也可以使用均匀标记的探针进行 S1 核酸酶分析，这使得 S1 核酸酶分析几乎与 RNase 保护一样灵敏。然而，杂交体仍然易发生呼吸，从而导致相当高的背景。在本书中对 S1 核酸酶方法将不做详细讨论，这是因为其使用有限。然而，S1 核酸酶方案在其他地方也可得到（如 Sambrook and Russell 2001，pp. 7.51-7.62)，并且来自本书第一版的详细信息，可向作者索取。

数据分析范例

在图 3.3C 所示的例子中，利用单链基因组 DNA 探针进行 S1 核酸酶实验，其中的单链基因组 DNA 探针相对于 TdT 转录起始位点（+1）从核苷酸－111 延伸到核苷酸+58。探针是从包含一段基因组 DNA 片段的质粒中制备的。首先用 *Bam*HI 在+58 处切割质粒，用小牛肠碱性磷酸酶脱去磷酸，用 [γ-^{32}P] dATP 和 T4 多核苷酸激酶进行放射性标记，然后用 *Sac*I 在－111 处切割，在非变性聚丙烯酰胺分离胶上分离单链放射性标记的探针。从凝胶中切下代表 5′ 端标记探针的短的放射性条带，再从凝胶上洗脱 DNA 并纯化，通过闪烁计数进行定量，然后与 mRNA 样品杂交。

图 3.3C 所示的结果来自以 mRNA 样品进行的反应，其中 mRNA 样品来自两个表达 TdT 基因的细胞系（泳道 2 和 3）和一个不表达该基因的细胞系（泳道 1)。重要的是要注意该图的这一面板来自比其他面板更长时间的放射自显影曝光，反映了使用 5′ 端标记探针进行 S1 核酸酶分析时灵敏度会降低。该结果就像 RNase 保护结果一样，在大部分凝胶中非常清晰。然而，在代表起始位点的区域有 4 个或 5 个强带，不像引物延伸或 RNase 保护中观察到的单一强带。迁移得最慢的强带定位于 TdT cDNA 的 5′ 端，与引物延伸或 RNase 保护所确定的结果相同。由于 RNA-DNA 双链末端的呼吸可能产生其他条带，结果使 S1 核酸酶产生相对高效的切割。这例证了以前形成的观点，即 RNA-DNA 双链不如 RNA-RNA 双链稳定，使得 S1 核酸酶程序稍微不及 RNase 保护法精确。

专题 3.4　核酸酶保护

S1 核酸酶程序（图 3.6）的概念基础与 RNase 保护程序的概念基础相似，除了放射性标记的探针通常是单链或双链的 DNA 分子，而不是 RNA 分子。通过多种程序可制备 DNA 探针，其中一个技术涉及利用 Klenow 片段和放射性标记的 dNTP 延伸与单链 M13 质粒退火的引物。该质粒应含有跨越预期转录起始位点的基因组 DNA 片段。放射性标记的反义探针在进行限制酶切割和凝胶纯化后，与 RNA 样品中的特定 mRNA 分子杂交，然后加入 S1 核酸酶消化 DNA 探针及退火后 mRNA 的单链区。通过变性聚丙烯酰胺凝胶电泳确定放射性标记的 S1 抗性产物的大小，最后进行放射自显影或磷屏成像分析。抗性产物的长度应相当于从探针的 5′ 端到转录起始位点的距离。

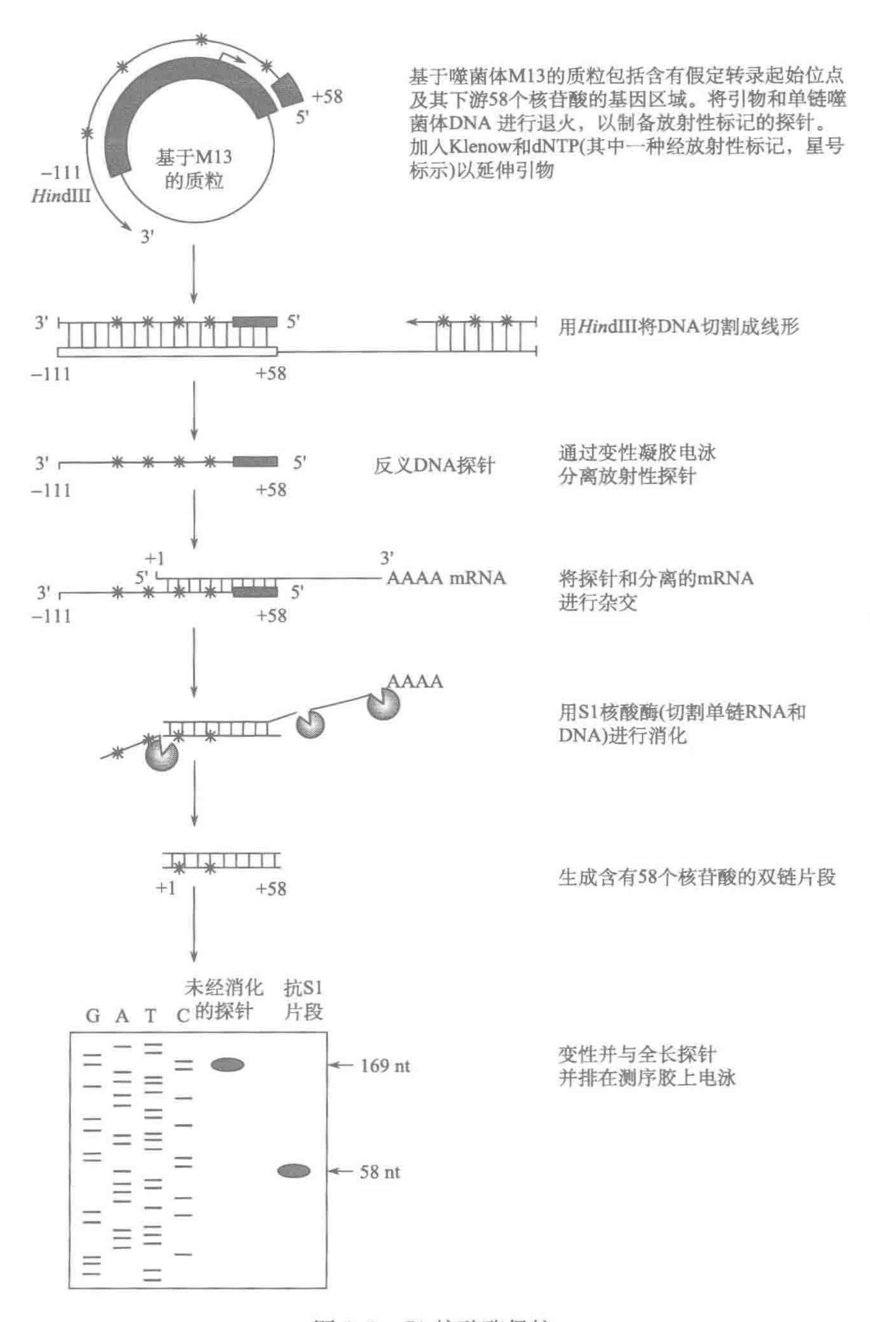


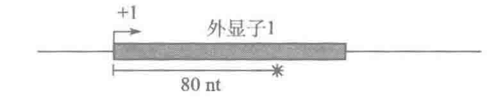
图 3.6　S1 核酸酶保护。

专题 3.5　内含子对起始位点定位结果解释的影响

在发展任何起始位点定位技术中一个重要的考虑是内含子对结果解释的影响（见图

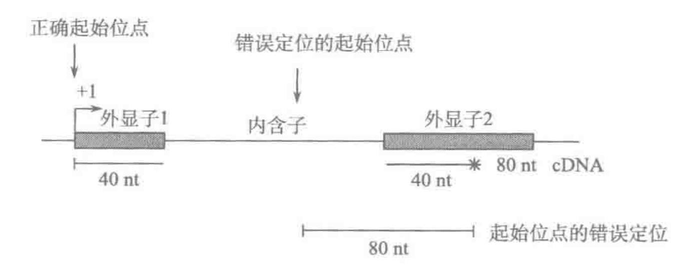
A

与外显子1的引物杂交



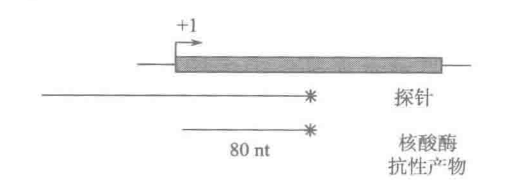
两种结构均会产生含80个核苷酸的cDNA 产物。如果引物和外显子2杂交（下图），但认为与外显子1杂交(上图)，就会错误地将起始位点指定为在第一个内含子中的一个位置(下图)。通过对经RACE获得的克隆进行测序并与基因组序列相比较，会消除这种错误

与外显子2的引物杂交



B

外显子1中探针的5' 端



两种结构均会产生含80个核苷酸的核酸酶抗性产物。如果探针跨越第一个内含子和第二个外显子之间的边界(下图)，而不跨越转录起始位点(上图)，则结果将指示探针到外显子-内含子连接处的距离(下图)，而非探针到转录起始位点的距离

外显子2中探针的5' 端

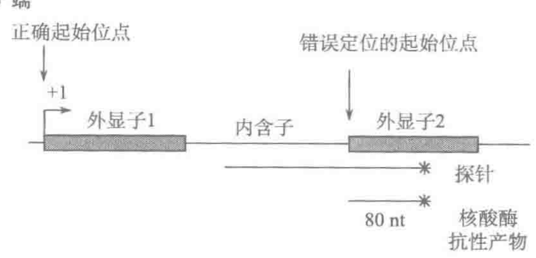


图 3.7 外显子-内含子结构对通过引物延伸分析和 RACE（A)，以及通过 RNase 保护和 S1 核酸酶分析（B）定位起始位点的影响。

3.7)。当进行引物延伸和 RACE 分析时，在引物结合位点和转录起始位点之间内含子的存在，可导致误导性的结果。同样，在 RNase 保护和 S1 核酸酶探针跨越的区域中，拼接受体或供体的存在也将导致错误解释的结果。必须强调内含子的潜在影响，因为研究者偶尔花费数月时间研究一个基于起始位点定位数据被假定包含启动子的 DNA 区域，才发现起始位点的位置已经被错误计算数百个到数千个碱基，因为一个内含子影响

了实验的结果。

对于引物延伸和RACE分析，cDNA的大小对应于从引物5′端到成熟mRNA的5′端的距离。如果引物与基因的第一外显子序列互补，且引物序列、基因组DNA序列、cDNA产物的大小已知，就能很容易地确定转录起始位点。如果引物与第二个外显子序列互补，依然能很容易地确定转录起始位点，但是必须已知第一个外显子的位置和序列及精确的拼接受体和供体位点。最重要的是，必须注意不能仅仅因为杂交信号在分离的cDNA 5′端旁边，就认为引物与第一个外显子杂交。如果认为引物与第一个外显子杂交，而实际上是与第二个外显子杂交，引物延伸程序会错误地提示起始位点位于第一个内含子中，这一概念如图3.7A所示。RACE程序具有防止错误解释数据的优点。当对来自RACE的单个克隆进行测序时，如果最初的引物与第二个外显子的序列杂交，那么得到的序列将与基因组DNA序列不共线。而且，RACE克隆的序列将偏离位于第二个外显子和第一个内含子边界处的基因组序列，从而反映内含子的存在。

对RNase保护和S1核酸酶方法，如果探针的5′端与第一个外显子内的一段序列杂交，对RNase或S1核酸酶切割具有抗性的探针片段的大小将对应于从标记探针的5′端到mRNA 5′端的距离。然而，如果探针的5′端与第二个外显子（或第一个内含子）内的一段序列杂交，则抗性产物将不会对应于所需的距离。获得的抗性产物可能对应于任一距离，这取决于探针杂交的确切部位及探针是内部标记还是5′端标记。内部标记的探针有可能产生一个条带，该条带对应于和mRNA的一段序列杂交的任何片段。相比之下，5′端标记的探针也会产生一个条带，但该条带对应于从标记核苷酸到外显子-内含子接合处的距离（如果标记的核苷酸杂交在第二个外显子中）。图3.7B说明了这些可能的问题。

技术

方案3.1　引物延伸分析

概论

20世纪70年代初期，随着分子生物学基本技术的发展及后来反转录酶的发现，诞生了利用反转录病毒反转录酶和DNA引物制备特定mRNA的cDNA的一般方法。几年后，该技术被用于转录起始位点的定位及相对mRNA浓度的确定（Ghosh et al. 1978；将专题3.2）。利用反转录酶的起点定位技术通常被称为引物延伸分析。

引物延伸分析需要特定的DNA引物，典型的DNA引物是约20个残基的人工合成寡核苷酸，该引物与目的mRNA 5′端附近的一段序列互补。用［γ-^{32}P］ATP和T4多核苷酸激酶对该引物进行5′端标记，并且引物与RNA样品中的特定mRNA分子退火。之后，利用反转录酶和未标记的核苷三磷酸将退火的引物分子延伸到mRNA 5′端。通过变性聚丙烯酰胺凝胶电泳及放射自显影确定放射性标记的cDNA产物的大小。理论上，cDNA的3′端将与mRNA的5′端一致。因此，放射性标记的cDNA大小应代表

从引物标记的 5′ 端到 mRNA 5′ 端的距离（即 cDNA 的 3′ 端）。例如，当与邻近 DNA 分子质量标准物比较时，如果 cDNA 产物在放射自显影照片中迁移至 50 个核苷酸大小处，那么转录起始位点可能位于距引物的标记 5′ 核苷酸 50 个核苷酸处。

可用从商业化来源购买的试剂盒（如 Promega 公司）进行引物延伸分析，或者按任一使用反转录酶的方案（如 Sambrook and Russell 2001，pp. 7.75-7.81）进行引物延伸分析，不同的反转录酶包括从大肠杆菌表达的反转录酶或从反转录病毒感染的细胞中纯化的反转录酶。下面是使用 MMLV 反转录酶的一种基本方案（改进自 McKnight and Kingsbury 1982）。与 RNase 保护分析不同，引物延伸分析很少得益于商业化试剂盒的使用。

时线与安排

第 1 天：引物延伸分析

第 2 天：数据分析

提纲

引物延伸分析（时间：1 天）

步骤 1：激酶处理纯化的寡核苷酸引物（2 h）。

步骤 2：除去未掺入的 [γ-^{32}P] ATP（30 min）。

步骤 3：配制缓冲液（30 min）。

步骤 4：引物延伸（3～4 h）。

步骤 5：聚丙烯酰胺凝胶电泳（2～3 h）

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

无核酸酶的缓冲液和试剂，以及核糖核酸酶抑制剂可用于 RNA 制备和引物延伸反应，但它们一般是不需要的。核糖核酸酶抑制剂需要充分降解，否则显著影响引物延伸实验的结果，因为对于引物延伸分析的成功，只要求每个 mRNA 分子 5′ 端是完整的。

放线菌素 D（可选，见步骤 4：9）〈!〉

在乙醇中准备 5 mg/ml 放线菌素 D 贮液，避光保存于－20℃。使用前立即加入到 RT 缓冲液中。此外，还可制备甘露醇溶解的放线菌素 D（Sigma 公司）溶液，并按上述方式存储。

10×噬菌体 T4 多聚核苷酸激酶缓冲液

50 mmol/L DTT（二硫苏糖醇）〈!〉

1 mmol/L EDTA（pH 8.0）

0.1 mol/L $MgCl_2$〈!〉

1 mmol/L 亚精胺 HCl〈!〉

0.5 mol/L Tris-HCl（pH 7.6）〈!〉

干冰〈!〉

乙醇〈!〉

[γ-^{32}P] ATP，高比活性（大于 5000 Ci/mmol；10 mCi/ml）〈!〉

MMLV 反转录酶（如 MMLV SuperScript III RT；Invitrogen）

5 mol/L NaCl

0.1 mol/L NaOH〈!〉

纯化的寡核苷酸

通常用纯化的合成寡核苷酸作为引物，并一般用 T4 多聚核苷酸激酶对引物进行有效标记。为了提高激酶反应的效率，应该通过凝胶或过柱等方法从合成的寡核苷酸制品中除去杂质（见 Sambrook and Russell 2001，pp. 10.48-10.49）。之后，测定稀释的引物样品的 OD_{260} 值，通过计算确定引物浓度（Sambrook and Russell 2001，pp. A8.20-8.21）。

5×引物延伸（PE）杂交缓冲液

5 mmol/L EDTA

1.25 mol/L KCl〈!〉

50 mmol/L Tris-HCl（pH 7.5）

反转录（RT）缓冲液

100 μg/ml BSA（牛血清白蛋白）

0.5 mmol/L dCTP、dGTP、dATP、dTTP

10 mmol/L DTT〈!〉

25 mmol/L KCl〈!〉

3.5 mmol/L $MgCl_2$〈!〉

50 mmol/L Tris-HCl（7.5）

保存于－20℃。

目的 RNA [总 RNA 或 poly（A）$^+$ mRNA]

可使用任一试剂盒或发表的方法（见 Sambrook and Russell 2001，pp. 7.4-7.26），从合适的细胞系或组织制备纯化的细胞质 RNA、细胞总 RNA 或多聚腺苷酸化的 mRNA。

3 mol/L 乙酸钠（pH 5.2）

STE 缓冲液

10 mmol/L EDTA

100 mmol/L NaCl

20 mmol/L Tris-HCl（pH 7.5）

T4 多聚核苷酸激酶（10 U/μl）

TE 缓冲液

设备

用于探针纯化的柱子（如 NucTrap；Stratagene）

蒸发器，SpeedVac
增感屏（可选）
离心机
离心管
磷屏成像仪（可选）
推式柱β屏蔽装置（Stratagene，400700）
闪烁计数器
BD 注射器，带 Luer 旋转锁（Luer-Lok）头（10 cm^3）
组织
X 光片（如 XAR-5 膜；可选）

程序

步骤 1：激酶处理纯化的寡核苷酸引物

注意：为提高激酶反应的效率，应通过凝胶或过柱等方法从人工合成的寡核苷酸制品中除去杂质（见 Sambrook and Russell 2001，pp. 10.48-10.49）。之后，测定稀释的引物样品的 OD_{260} 值，通过计算确定引物浓度（Sambrook and Russell 2001，pp. A8.20-8.21）。

1. 混合 5 μl 10×T4 多核苷酸激酶缓冲液、2.5 pmol 凝胶纯化的寡核苷酸及 5 μl（约 10pmol）[γ-^{32}P] ATP，加 dH_2O 至终体积 49 μl。
2. 加 1μl（10 U）T4 多核苷酸激酶。
3. 37℃孵育 45～60 min。
4. 68℃孵育 10 min，热失活激酶。

步骤 2：除去未掺入的 [γ-^{32}P] ATP

注意：应从制备的引物中除去未掺入的 [γ-^{32}P] ATP，以防止它们干扰引物延伸反应，特别是凝胶电泳和放射自显影。可以通过在乙酸铵存在的情况下进行示差乙醇沉淀或用柱层析法除去未掺入的核苷酸（Sambrook and Russell 2001，pp. 10.20-10.21）。

1. 向热失活反应样品中加入 1 μl 5 mol/L NaCl。
2. 向 Stratagene NucTrap 柱加入 70 μl STE 缓冲液，预洗柱子。
3. 按生产商的方案，用 Luer-Lok 头延长带活塞的 10 cm^3 BD 注射器，将注射器旋转连接到柱上，使柱和注射器间密封，同时注意不要将注射器拧得太紧。
4. 用力推进注射器活塞，使缓冲液向下流经柱子，直至柱的末端存在只有一小滴缓冲液。该步骤要在 5～10 min 内使用预洗过的柱子完成。
5. 按生产商的操作指南将柱安装到β屏蔽装置上。
6. 将放射性溶液加到柱上，连上延长的注射器，缓慢（25～35s）推进活塞，使样品流经柱子。用微量离心管收集过柱液，未掺入的核苷酸将留在树脂中，放射性标记的引物流经柱子进入收集管。

7. 将 70 μl STE 溶液加到柱上，用注射器推过柱子，将过柱液收集在同一收集管中。

8. 用另外的 140 μl STE 缓冲液稀释引物溶液。

9. 通过闪烁计数（切伦科夫计数；即在闪烁液存在的条件下测定放射活性），虽然切伦科夫测量因使用的机器不同而不同，但在最终的引物溶液中应获得大致 50 000～100 000 cpm/μl 放射活性。

10. 将引物贮存在−20℃。该引物溶液通常可以使用长达 1 个月。

步骤 3：引物退火和延伸缓冲液的配制

1. 配制 5×PE（引物延伸）杂交缓冲液和 RT（反转录）缓冲液（−20℃保存）。

步骤 4：引物延伸

注意：每个引物的最佳杂交温度应该依经验确定。一个确定最佳温度的好的起点是测试下列温度：37℃、45℃、60℃和 68℃。有些引物在杂交反应中先加热到 68℃，然后缓慢降温到室温时效果较好。

应尽可能设立阴性对照 RNA，这对定位未知起始位点很重要。tRNA 或酵母 RNA 对照产生的信息很少，但可用作附加对照。最理想的是，应从不表达测定 RNA 的细胞系或组织中制备 RNA 作为阴性对照，并且这个细胞系最好来自作为实验用 RNA 的同一物种。如果可能，应测定从多个阳性和阴性组织或细胞制备的 RNA。如果没有合适的阴性对照，一个新的起始位点的定位的说服力不强。

1. 将 1 μl 放射性标记的引物（50 000～100 000 cpm，由切伦科夫计数测定）加入 10～60 μg总 RNA［或 2～5 μg poly (A)$^+$ mRNA］中。引物延伸反应的最佳 RNA 用量会有变化，可能需要依经验而定。一般来说，该程序的成功与 RNA 的用量关系不大，而与引物的比活性、退火温度、RNA 二级结构的有无等变量关系密切。在最初的实验中，30 μg 总细胞 RNA 或总细胞质 RNA 是适宜用量。

2. 加入 3 mol/L 乙酸钠（pH5.2）使终浓度为 0.3 mol/L，再加入 2.5 倍体积的乙醇。在干冰上放置 10 min 或在−20℃放置 30 min 以沉淀 RNA 和引物。

3. 14 000 *g* 离心 10 min，沉淀 RNA 和引物。

4. 小心弃去上清，在 Speedvac 旋转蒸发仪中干燥沉淀或在台面上风干沉淀，放射性应主要存在于沉淀中，在上清中很少。

5. 将干燥的沉淀悬浮于 8 μl 的 TE 缓冲液中。一般通过涡旋难以将沉淀悬浮于这么少的溶液中，因为缓冲液不会保留在管的底部。悬浮沉淀的首选方法是用手指反复轻弹管底。或者可以用移液器反复吹打溶液使沉淀溶解。

6. 向管中加入 2 μl 5×PE 缓冲液。混合均匀并简单离心。

注意：只有在 RNA 完全溶解于 TE 溶液中后才能加 5×PE 溶液，因为 RNA 在高盐 PE 缓冲液中溶解很慢。

7. 在依经验确定的温度下孵育 90 min，使引物与 RNA 退火。在孵育期间，每隔 20 min 需要对样品简单离心，以防止溶液凝结在管的顶部。

8. 为每个反应配制一份含有 39.5 μl RT 缓冲液和 0.5 μl MMLV 反转录酶的溶液。或者向 RT 缓冲液中加入 50 μg/ml 的放线菌素 D，以抑制反转录酶合成双链 DNA（见疑难解答）。由于移液器不是非常准确，所配溶液量最好比实际需要量要多一些。

注意：我们发现来自 Invitrogen 公司的克隆的 MMLV Superscript III RT 效果好（cat. ＃18080-093，18080-044，或 18080-085），但使用其他来源的酶也能得到非常好的结果。SuperScript III 反转录酶含有突变，该突变能增加酶的热稳定性并可使延伸反应在较高温度下完成，这使其潜力减少或消除由 RNA 二级结构引起的问题。

9. 向每个反应管中加入 40 μl RT 混合液。混匀，不要旋涡。简单离心。

10. 37℃孵育 60 min。可以尝试用较高的温度（如 42℃、45℃、55℃），以降低 RNA 二级结构的不良影响。

11. 加入 6 μl 3 mol/L 乙酸钠和 150 μl 乙醇，终止反应并沉淀核酸。充分涡旋，在干冰上放置 10 min 或在－20℃放置 30 min。

12. 离心 10 min。可以在管底看见少量白色沉淀。弃去乙醇，将管倒置在纸巾上除去多余的液体。大部分放射性应留在沉淀中。

13. 用 SpeedVac 旋转蒸发仪充分干燥沉淀。

步骤 5：变性聚丙烯酰胺凝胶电泳

1. 将沉淀悬浮在 4 μl 甲酰胺上样染料及 2 μl 0.1 mol/L NaOH 中。

注意：此时，每管中的放射性量应仍然与在步骤 4：1 中加入反应的量相似，如果量减少，很可能是沉淀效果不好或沉淀损失。

2. 煮沸 2 min 使 DNA/RNA 杂交体变性。简单离心。

3. 冲洗凝胶点样孔后，每个样品上样 3 μl 到 8%变性聚丙烯酰胺测序胶上。

4. 跑胶，直到溴酚蓝接近凝胶底部（目的是让过剩的标记引物迁移到接近凝胶底部）。

步骤 6：产物分析

1. 干燥凝胶，并在增感屏存在的情况下在 XAR-5 膜上曝光凝胶，或在磷屏成像仪上分析。曝光过夜后膜上的强信号应该是可检测的。

附加说明

1. 上述方案是为 MMLV 反转录酶设计的。其他的反转录酶，如禽类成髓细胞瘤病毒（AMV）反转录酶也普遍应用，但需要不同的反应缓冲液。在第 12 章中可找到一个使用 AMV 反转录酶的方案，该方案可与体外转录方案联用。

2. 无核酸酶的缓冲液和试剂，以及核糖核酸酶抑制剂可用于 RNA 制备和引物延伸反应，但它们一般是不需要的。核糖核酸酶抑制剂需要充分降解，否则会显著影响引物延伸实验的结果，因为对于引物延伸分析的成功，只要求每个 mRNA 分子的 5′ 端是完整的。

3. 放线菌素 D 抑制反转录酶合成双链 DNA，并且有时用在反转录酶反应中，以抑

制发夹分子的合成。使用人工合成的短寡核苷酸引物时，发夹结构一般不会造成问题，除非预期的延伸产物也很短。如果需要，可以将放线菌素 D 以 50 μg/ml 的浓度加到反转录缓冲液中。放线菌素 D 的贮液是 5 mg/ml 的乙醇溶液，避光保存于－20℃，使用前立即加到缓冲液中。还可以配制放线菌素 D 的甘露醇溶液（Sigma cat. ＃ A-5156），保存条件同上。

4. 通常用 T4 多核苷酸激酶对人工合成的寡核苷酸进行有效的标记。因此，通常不需要对掺入的放射性进行精确定量。正如步骤 2 中所描述的，对放射性标记的引物进行最后的切伦科夫计数通常就足够了。如果担心激酶处理的效率，可以使用更严格的定量掺入的方法，如 Sambrook 和 Russell 所述的方法（Sambrook and Russell 2001，pp. 8.25-8.26）。

疑难解答

引物延伸分析的疑难很难解答，因为每个问题有多种潜在的原因，并且每个潜在的原因都有多个可能的解决办法。最普遍的问题是信号弱和条带多，从而很难鉴别哪条带对应于正确的起始位点。如果已付出了相当大的精力仍不能获得清楚而令人信服的转录起始位点，有必要采用至少一种其他分析方法（即 5′ RACE、RNase 保护、S1 核酸酶分析），以增强发现该起点位置的信心。即使采用多种分析方法，有时也难以信心满满地确定一个基因的转录起始位点。

放射性标记引物产量低

可能原因：试剂（如激酶可能已经失活、放射性标记的 ATP 可能被水解、激酶缓冲液的配制不准确或 DTT 可能被氧化）。

解决办法：测试新的试剂，采用更严格的定量 ^{32}P 掺入的方法（Sambrook and Russell 2001，pp. 8.25-8.26）。

可能原因：寡核苷酸可能含有过多的杂质或定量不精确。

解决办法：纯化新的一等份寡核苷酸或确证浓度。将产量与使用另一个寡核苷酸获得的产量进行比较。利用不同的方法纯化寡核苷酸，采用更严格的定量 ^{32}P 掺入的方法（Sambrook and Russell 2001，pp. 8.25-8.26）。

可能原因：除去未掺入 ATP 的方法可能导致放射性标记引物的损失。

解决办法：重复程序。试用不同的方法去除未掺入的核苷酸。

大部分放射性存在于上清液中

步骤 4：4 开始时，乙醇沉淀后大部分放射性保留在上清液中。

可能原因：在步骤 2：6 中，去除未掺入的 ATP 的效率太低。

解决办法：用剩余的放射性标记引物制剂重复步骤 2：6。试用不同方法去除未掺入的核苷酸。

可能原因：用乙醇沉淀放射性标记引物的效率太低。

解决办法：在冷冻处理前更充分地混合乙醇/引物/RNA 溶液。在干冰上冷却更长

的时间。加入另一份 0.5 倍体积的乙醇和 0.1 mol/L 的乙酸钠。

可能原因：可能将沉淀带入了上清液中。

解决办法：再次离心上清液，在去除上清时设法检测沉淀。用拉长的巴斯德吸管除去上清液。

空白凝胶图像、无引物信号、无 cDNA 产物

可能原因：放射性标记引物产量可能已经很低。引物制剂中的放射性可能已经主要归因于未掺入的核苷酸。

解决办法：用另一份放射性标记引物进行凝胶电泳，并将凝胶在膜上曝光以确定制剂是否含有放射性标记引物。如果没有，按照“疑难解答：放射性标记引物产量低”的方法进行。

可能原因：放射性标记的引物可能在乙醇沉淀步骤中丢失。

解决办法：在每次沉淀过程中仔细测定上清液和沉淀中放射性的量。

延伸产物太弱

凝胶图像表明在凝胶的底部有强的引物信号，但没有延伸产物，或只有短延伸产物，并且这些短延伸产物在阴性对照中也出现。

可能原因：与特定 mRNA 退火效率低。

解决办法：试用不同的退火温度，试用不同的退火缓冲液组分，如甲酰胺（见 Sambrook and Russell 2001，p. 7.67）。试用与稍远端或稍近端序列退火的其他不同引物。在某些情况下，需要测定多个不同引物以确定能发生有效退火的引物。某些引物退火效率低的原因尚不清楚。

可能原因：反转录酶的延伸效率低。

解决办法：利用较高浓度或不同批次的反转录酶。制备一批新的反转录酶缓冲液。在较高温度下进行反转录反应以减少 mRNA 的二级结构。检测反应缓冲液的 pH（特别是对 pH 敏感的禽类反转录酶）。

可能原因：特定 mRNA 浓度太低，无法检测。

解决办法：用更多的 mRNA 进行实验。试用纯化的多聚腺苷酸化 mRNA。试用可能更灵敏的定位转录起始位点的方法，如 RNase 保护方法或 RACE。

凝胶图像上背景带太多

难以区分特异性 cDNA 产物与背景。

可能原因：引物与样品中其他 RNA 杂交。

解决办法：试用较高的退火温度和（或）延伸温度。试用与稍近端或稍远端序列互补的其他引物。利用不同的方法，如 RNase 保护、S1 核酸酶分析、RACE 确证起始位点。减少反转录酶用量。

可能原因：反转录酶的延伸效率太低。

解决办法：试用较高的延伸温度来减少 RNA 的二级结构。向反转录缓冲液中加入

放线菌素 D。加入更多的反转录酶或试用新的反转录缓冲液。试用可能包含较少反转录酶抑制性杂质的其他 RNA 制剂。

方案 3.2　RNase 保护分析

概论

RNase 保护分析在 20 世纪 80 年代初期产生，作为灵敏的方法用于定量在特定核苷酸处起始的 mRNA 转录本（Zinn et al. 1983），后来证明这种方法对于定位转录起始位点很有用，并用于研究 RNA 的 3′ 端及拼接连接区。

为了确定转录起始位点，必须制备跨越预期起始位点的放射性标记 RNA 探针。可以通过将基因组 DNA 片段以反义方向插入在几种市售载体之一中的噬菌体 T7、T3 或 SP6 启动子的下游。或者用相关的基因组 DNA 片段进行 PCR 反应来制备探针，其中一个引物包含噬菌体启动子序列。利用噬菌体 RNA 聚合酶从基因组片段制备反义转录本，并在反转录过程中用［α-^{32}P］NTP 进行全转录本的高比活放射性标记。分离标记的探针，有时凝胶纯化，并与 RNA 样品中的特定 mRNA 分子退火。之后，用 RNase 消化 RNA 中没有退火（单链）的区域。理论上，没有被消化的放射性标记 RNA 片段的大小相当于从目的 mRNA 转录起始位点到包含在探针内的基因组片段末端的距离（见专题 3.3）。RNase 保护探针的理想长度难以确定。较长的探针（250 个核苷酸或以上）具有较高的比活性优势，因为大量放射性标记的核苷酸可掺入到每个探针分子中。此外，短探针（少于 250 个核苷酸）可能是有利的，因为使用噬菌体 RNA 聚合酶提高了全长探针的生产效率，且探针的凝胶纯化往往没有必要。

强烈建议为 RNase 保护分析的初始建立使用商品试剂盒，这些试剂盒一般包括精确滴定的试剂和有用的对照模板。从 Zinn 等（1983）及 Sambrook 和 Russell（2001，pp. 7.63-7.74）改编的一个替代程序如下。正如引物延伸分析，RNase 保护分析最佳的阴性对照是来自于不期望包含目的基因的组织或细胞的 RNA。如果可能，RNA 应来自与实验中所使用的 RNA 相同的物种。制备好线性质粒模板（或由 PCR 产生的模板）及细胞 RNA 之后，一般可在约两天内完成 RNase 保护分析。

时线与安排

在开始 RNase 保护分析前，必须将跨越预期转录起始位点的基因组 DNA 片段以反义方向克隆到含有噬菌体 T7、T3 或 SP6 启动子的合适载体中（可从 Promega、Stratagene 或 Ambion 获得），也可以利用含有噬菌体启动子序列的引物进行 PCR 扩增。在利用质粒制备 RNA 探针之前，必须用合适的限制酶进行切割，这将使噬菌体转录在特定位点终止，产生大小不一的放射性标记探针分子（见专题 3.3）。限制酶酶切应该产生平末端或 5′ 凸出末端。酶切后的质粒 DNA（或 PCR 产物）应以酚：氯仿抽提，并用乙醇沉淀（Sambrook and Russell 2001，pp. A8.9-A8.10，A8.12-A8.16）。在进行 RNase 保护分析之前，还需要从合适的细胞系或组织中纯化细胞质 RNA、细胞总 RNA 或多聚腺苷酸化的 mRNA。可以用任一试剂盒或已发表的方法纯化 RNA（见 Sam-

brook and Russell 2001，第 7 章）。制备好线性质粒模板（或由 PCR 获得的模板）及细胞 RNA 之后，一般可在两天内完成 RNase 保护分析。

第 1 天：制备放射性标记的 RNA 探针并与特定 mRNA 退火。

第 2 天：RNase 消化和凝胶电泳。

第 3 天：数据分析。

提纲

制备放射性标记的 RNA 探针并与特定 RNA 退火（1～2 天）。

步骤 1：配制缓冲液（1 h）。

步骤 2：制备探针（2 h）。

步骤 3：通过 PAGE 纯化探针（可选，2～3 h，之后温育过夜）。

步骤 4：探针与 RNA 杂交（1 h，之后温育过夜）。RNase 消化和凝胶电泳（时间投入：1 天）。

步骤 5：利用 RNase 消化未退火的 RNA（1 h，30 min）。

步骤 6：聚丙烯酰胺凝胶电泳（2～4 h）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

为了减少 RNase 的污染，可用 DEPC（焦碳酸二乙酯）处理的水配制溶液。

[α-^{32}P] UTP（10 mCi/ml；800 Ci/mmol）〈!〉

可以使用其他放射性标记的核苷酸。对一些噬菌体聚合酶而言，GTP 具有较低的 K_m 值，这允许在探针制备过程中可使用更高比活性的放射性核苷酸。这样可制得更高比活性的探针。ATP 的 K_m 值特别高，很少用于探针的制备。

2.5 mol/L 乙酸铵

ATP、CTP、GTP（各 10 mmol/L）

氯仿〈!〉

线性 DNA 模板（1 mg/ml）

用限制性内切核酸酶裂解质粒，该酶将使噬菌体转录在一特定位点终止，从而产生不同大小的放射性标记探针分子。限制性内切核酸酶产生一个平末端或 5′ 凸出末端。使用前对酶切后的质粒 DNA（或 PCR 产物）应使用酚：氯仿抽提，并用乙醇沉淀（Sambrook and Russell 2001，pp. A8.9-A8.10，A8.12-A8.16）。

DNase I（无 RNase；2 U/μl）

干冰〈!〉

100 mmol/L DTT〈!〉

荧光染料（可选，见步骤 8）

洗脱缓冲液（可选，见步骤第 8）

0.5 mol/L 乙酸铵
1 mmol/L EDTA
0.1 %SDS（十二烷基硫酸钠）〈!〉
乙醇〈!〉
甲酰胺凝胶上样缓冲液
0.05% 溴酚蓝〈!〉
20 mmol/L EDTA
95% 甲酰胺〈!〉
0.05% 二甲苯青〈!〉
杂交缓冲液：
1 mmol/L EDTA
80% 甲酰胺〈!〉
400 mmol/L NaCl
40 mmol/L PIPES（哌嗪-N，N′-二（2-乙磺酸））（pH6.4）
苯酚〈!〉
蛋白酶 K（10 mg/ml）〈!〉
目的 RNA [总 RNA 或 poly（A）$^+$ mRNA]

可以用任一试剂盒或已发表的方法从合适的细胞系或组织制备纯化的细胞质 RNA、细胞总 RNA 或多聚腺苷酸化的 mRNA（Sambrook and Russell 2001）。

RNA 聚合酶（SP6、T7 或 T3；5 U/μl）
RNase 消化缓冲液：
5 mmol/L EDTA
300 mmol/L NaCl
10 mmol/L Tris-HCl（pH 7.5）
临用前，加入
40 μg/ml RNase A〈!〉
2 μg/ml RNase T1
胎盘 RNase 抑制剂（10 U/μl）
20% SDS〈!〉
0.3 mol/L 乙酸钠
5×转录缓冲液
30 mmol/L $MgCl_2$〈!〉
50 mmol/L NaCl
10 mmol/L 亚精胺〈!〉
200 mmol/L Tris（pH 7.5）〈!〉
酵母 RNA（或 tRNA）

设备

培养箱，37℃
增感屏（可选）
离心机
离心管
磷屏成像仪（可选）
保鲜膜（可选）
手术刀或剃刀刀片（可选）
闪烁计数器
水浴
X 光片（可选）

程序

步骤 1：配制缓冲液

配制以下缓冲液：5×转录缓冲液、洗脱缓冲液（如果探针需要凝胶纯化，则需要配制）、杂交缓冲液和 RNase 消化缓冲液。

步骤 2：制备探针

1. 在一离心管中混合下述组分。在室温下按如下顺序混合：

dH_2O	2μl
线性模板 DNA（1mg/ml）	1 μl
10 mmol/L ATP	1 μl
10 mmol/L CTP	1 μl
10 mmol/L GTP	1 μl
100 mmol/L DTT	1 μl
5×转录缓冲液	4 μl
胎盘 RNase 抑制剂（10 U/μl）	1 μl
[α-^{32}P] UTP（10 mCi/ml；800 Ci/mmol）	5 μl
SP6、T7 或 T3 RNA 聚合酶（5 U/μl）	2 μl
总体积	20 μl

2. 轻弹管壁将溶液混匀（不要旋涡），然后简单离心。
3. 将反应混合物于 37℃孵育 1 h。
4. 向反应中加入 1 μl 不含 RNase 的 DNase I（2 U/μl）。消化模板 DNA，轻弹管壁混匀并简单离心。
5. 将反应置于 37℃孵育 15 min。

注意：有些探针，特别是较长的探针（大于 100 个核苷酸），可能需要凝胶纯化以去除较短的 RNA 产物，这些产物会增加在 RNase 保护分析中观察到的背景产物的数量。放射性标记探针可以用凝胶电泳纯化（参见步骤 3：用 PAGE 纯化探针，见下文）。

6. 向反应中加入 80 μl 2.5 mol/L 乙酸铵和 5 μg 酵母 RNA（或 tRNA）。

7. 用 100 μl 酚∶氯仿抽提，然后用 100 μl 氯仿抽提。

8. 加入 300 μl 乙醇，在干冰上冷却 10 min 或在－20℃冷却 30 min 。

9. 在微量离心机中于 4℃离心 10 min ，以沉淀 RNA。

10. 用 100 μl 0.3 mol/L 乙酸钠重悬 RNA，再加入 300 μl 乙醇，按上面的方法沉淀。

11. 干燥 RNA 沉淀。

12. 将沉淀悬浮于 50μl 的杂交缓冲液中。

13. 取 1μl 在闪烁计数器中测定探针的活性。

14. 探针可在－20℃保存数天，但当新制备的探针用于随后的步骤时，可取得最佳的效果。

步骤 3：通过 PAGE 纯化探针

按步骤 2：5 进行凝胶纯化。

1. 在步骤 2：4 和步骤 2：5 DNase 消化后，向反应混合液中加入等体积的甲酰胺凝胶上样缓冲液（21 μl；95%甲酰胺、20 mmol/L EDTA、0.05%溴酚蓝、0.05%二甲苯青），并在 85～95℃加热 3～5 min。

2. 将整个反应上样到 8%的变性聚丙烯酰胺凝胶上，电泳。

3. 从凝胶上切下全长的片段，用 350 μl 洗脱缓冲液洗脱过夜。

4. 通过离心使凝胶片段与探针分开，然后将含有放射性 RNA 的水溶液转移到新的离心管中。可对凝胶再次洗脱、离心并转移水溶液，用于从凝胶片段中抽提剩下的探针。

5. 在闪烁计数器中计数 1μl 样品，确定探针的比活性。

6. 探针在洗脱液中－20℃可保存 3～4 天。

步骤 4：探针与 RNA 杂交

1. 将标记的探针［（2～8）×10^4 cpm］与 10～30 μg 的每个 RNA 样品混合。

2. 加入 0.1 倍体积的 3 mol/L 乙酸钠（pH5.2）和 2.5 倍体积的乙醇。充分混合并置于干冰上 10 min。

3. 4℃下 14 000 r/min 离心 15 min 沉淀探针和 RNA。

4. 小心弃除各管中的乙醇，在台面上室温干燥沉淀物。

5. 通过涡旋将沉淀重悬于 30 μl 的杂交缓冲液中，然后简单离心。

6. 将反应物于 90℃加热 3～4min，涡旋混匀并简单离心。

7. 立即将管置于水浴中；水浴温度设置在依经验而定的退火温度（通常 45℃），并在此步前预热在此温度下。

8. 杂交过夜（或 8 h）。

步骤 5：用 RNase 消化杂交探针

1. 向每个反应中加入 300 μl RNase 消化缓冲液（含有 RNase），混匀并简单离心。

2. 各反应管于 30℃孵育 60 min。

3. 加入蛋白酶 K（10 μl 的 10 mg/ml 溶液）和 SDS（10 μl 的 20%的溶液）终止 RNase 消化反应，混合并简单离心，于 37℃孵育 15min。

4. 用 350 μl 酚：氯仿抽提。

5. 加入 2 μg tRNA 或酵母 RNA 和 1 ml 乙醇，在干冰上冷却 10 min 或在－20℃冷却 30 min。

6. 4℃下 14 000 r/min 离心 15 min 沉淀 RNA。

7. 小心除去各管中的所有上清液，残余的上清液将导致条带在凝胶中的异常迁移。

8. 在台面上室温下干燥沉淀。

9. 将沉淀溶解于 8 μl 甲酰胺凝胶上样缓冲液中。

步骤 6：凝胶电泳和产物分析

1. 将各管于 90℃处理 3～4 min。

2. 涡旋混匀并简单离心，将样品加到 8%聚丙烯酰胺凝胶上电泳，直至溴酚蓝迁移至凝胶底部。

3. 将凝胶干燥，并在带增感屏的 X 光片上曝光或用磷屏成像仪分析。

附加说明

1. 很难确定 RNase 保护分析中探针的理想长度。较长的探针（≥250 个核苷酸）具有较高的比活性，这是因为大量的放射性标记核苷酸可以掺入到每个探针分子中。短的探针（＜250 个核苷酸）也有好处，因为噬菌体 RNA 聚合酶产生全长探针的效率会提高，并且通常不需要再对探针进行凝胶电泳纯化。

2. 上面制备探针的方案利用放射性标记的 UTP，也可以用其他放射性标记的核苷酸代替 UTP。对于某些噬菌体聚合酶，GTP 有较低的 K_m 值，在探针制备过程中可以利用高比活的放射性寡核苷酸，产生具有更高比活性的探针。而 ATP 的 K_m 值特别高，因此很少用于制备探针。

3. 与引物延伸分析一样，RNase 保护分析中的最佳阴性对照是来自不含目的基因的细胞系或组织中的 RNA。如果可能，阴性对照 RNA 应该与用于实验的 RNA 来源于相同物种。

疑难解答

探针产量低

可能原因：聚合酶在偶然位点上终止，从而使 RNA 合成的效率较低。

解决办法：增加噬菌体 RNA 聚合酶和（或）放射性标记核苷酸的浓度，对产生的全长探针进行凝胶电泳纯化（尽管大多数转录本可能仍然会提前终止，但仍有可能获得足量的全长探针用于后面的实验）。将相同的基因组片段插入到不同噬菌体聚合酶的启动子下游，或试用仍然跨越预期转录起始位点的稍短的基因组片段。

可能原因：试剂或试管有 RNase 污染。

解决办法：遵照消除和抑制 RNase 的标准程序（Sambrook and Russell 2001，pp. 7.82-7.84）。使用高度纯化的质粒 DNA。

可能原因：试剂没有活性或配制不当。

解决办法：试用新批次的试剂，特别是放射性标记核苷酸、噬菌体 RNA 聚合酶和模板 DNA。利用高度纯化的质粒 DNA 来降低 RNase 的污染。

凝胶图像上没有 RNase 抗性产物

可能原因：RNase 对探针/RNA 杂交体过度消化。

解决办法：降低 RNase 浓度或减少消化时间，试用不同浓度的 RNase A 和 RNase T1。在没有 RNase A 的条件下试用 RNase T1。在探针/RNA 杂交体中不加 RNase，作为对照。用该对照应该可以观察到全长探针。

可能原因：探针和特定 RNA 退火效率低。

解决办法：在甲酰胺杂交缓冲液中，多数探针与特定 mRNA 在 45℃退火。然而，有些探针的最佳杂交温度可能有所不同。

可能原因：试剂或试管污染 RNase。

解决办法：遵照消除和抑制 RNase 的标准程序（Sambrook and Russell 2001，pp. 7.82-7.84）。使用高度纯化的质粒 DNA。

可能原因：特定 mRNA 量低或缺乏。

解决办法：实验中使用较高浓度的细胞 RNA。使用以寡聚（dT）层析法筛选的多聚腺苷酸化 mRNA。制备较长的探针来增加比活性。

凝胶图像上出现多重 RNase 抗性产物

可能原因：RNase 对探针/RNA 杂交体的消化不充分。

解决办法：增加 RNase 的浓度或延长消化时间。试用不同浓度的 RNase A 和 RNase T1。

可能原因：探针 RNA 自身退火区、与模板的对立链产生的探针分子退火，或与细胞 RNA 样品中其他 RNA 退火。

解决办法：进行对照实验来检测这些可能性。在反应中没有 mRNA 时，自身退火或与对立链产生的探针分子退火将产生 RNase 抗性产物。当实验使用来自不表达目的基因的细胞系的 RNA 时，与细胞 RNA 样品中其他 RNA 分子退火可能产生 RNA 抗性产物。RNA 应来自相同物种。通过设置适当的对照，可能就没有必要去除不需要的条带，因为可鉴定出不是来自特异性杂交的条带。然而，较短的探针、较高的杂交温度或增加 RNase 的浓度都有助于消除这些不需要的条带。

参考文献

Bajic V. B., Brent M. R., Brown R. H., Frankish A., Harrow J. Ohler U., Solovyev V. V., and Tan S. L. 2006. Performance assessment of promoter predictions on ENCODE regions in the EGASP experiment. *Genome Biol*. (suppl. 1) 17: S3. 1-S3. 13.

Barrera L. O., Li Z., Smith A. D., Arden K. C., Cavenee W. K., Zhang M. Q., Green R. D., and Ren B. 2007. Genome-wide mapping and analysis of active promoters in mouse embryonic stem cells and adult organs. *Genome Res*. 18: 46-59.

Berk A. J. and Sharp P. A. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* 12: 721-732.

Bock C., Walter I., Paulsen M., and Lengauer I. 2007. CpG island mapping by epigenome prediction. *PLuS Comput. Biol*. 3: e110. doi: 10. 1371/journal. pcbi. 0030110.

Carninci P., Kvam C., Kitamura A., Ohsurni T., Okazaki Y., Itoh M., Kamiya M., Shibata K., Sasaki N., Izawa M., et al. 1996. High-efficiency full-length cDNA cloning by biotinylated CAP trapper. *Genomics* 37: 327-336.

CLONTECHniques. 1996. CapFinder PCR cDNA Library Construction Kit. *CLONTECHniques XL*: 2-4.

Davuluri R. V., Grosse I., and Zhang M. Q. 2001. Computational identification of promoters and first exons in the human genome. *Nat. Genet*. 29: 412-417.

Edery I., Chu L. L., Sonenberg N., and Pelletier J. 1995. An efficient strategy to isolate full-length cDNAs based on an mRNA cap retention procedure (CAPture). *Mol. Cell. Biol*. 15: 3363-3371.

Frohman M. A. 1995. Rapid amplification of cDNA ends. In *PCR primer: A laboratory manual* (ed. C. W. Dieffenbach and G. S. Dveksler), pp. 381-409. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Frohman M. A., Dush M. K., and Martin G. R. 1988. Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci*. 85: 8998-9002.

Gardiner-Garden M. and Frominer M. 1987. CpG islands in invertebrate genomes. *J. Mol. Biol*. 196: 261-282.

Ghosh P. K., Reddy V. B., Swinscoe J., Lebowitz P., and Weissman S. M. 1978. Heterogeneity and 5′-terminal structures of the late RNAs of simian virus 40. *J. Mol. Biol*. 126: 813-846.

Hash imoto S., Suzuki Y., Kasai Y., Morohoshi K., Yamada T., Sese J., Morishita S., Sugano S., and Matsushima K. 2004. 5′-end SAGE for the analysis of transcriptional start sites. *Nat. Biotech*. 22: 1146-1149.

Ioshikhes I. P. and Zhang M. Q. 2000. Large-scale human promoter mapping using CpG islands. *Nat. Genet*. 26: 61-63.

Kasai Y., Hashimoto S., Yamada T., Sese J., Sugano S., Matsushima K., and Morishita S. 2005. 5′SAGE: 5-end Serial Analysis of Gene Expression database. *Nucleic Acids Res*. 33: D550-D552.

Kim T. H., Barrera L. O., Zheng M., Qu C., Singer M. A., Richmond T. A., Wu Y., Green R. D., and Ren B. 2005. A high-resolution map of active promoters in the human genome. *Narure* 436: 876-880.

Kodzius R., Kojima M., Nishiyori H., Nakamura M., Fukuda S., Tagami M., Sasaki D., Imamura K., Kai C., Harbers M., et al. 2006. CAGE: Cap analysis of gene expression. *Nat. Methods* 3: 211-222.

Kozak M. 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Ret*. 15: 8125-8148.

Kozak M. 1996. Interpreting cDNA sequences: Some insights from studies on translation. *Mamm. Genome* 7: 563-574.

Lo K., Landau N. R., and Smale S. T. 1991. LyF-1, a transcriptional regulator that interacts with a novel class of promoters for lymphocyte-specific genes. *Mol. Cell. Biol*. 11: 5229-5243.

Lockard R. E. and RajBhandary U. L. 1976. Nucleotide sequences at the 5 termini of rabbit α and β globin mRNA. *Cell* 9: 747-760.

Loh E. Y., Elliot J. F., Cwirla S., Lanier L. L., and Davis M. M. 1989. Polymerase chain reaction with single-sided specificity: Analysis of T cell receptor chain. *Science* 243: 217-220.

Lynch M., Scofield D. G., and Hong X. 2005. The evolution of transcription-initiation sites. *Mol. Biol. Evol.* 22: 1137-1146.

Maruyama K. and Sugano S. 1994. Oligo-capping: A simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* 138: 171-174.

McKnight S. L. and Kingsbury R. 1982. Transcriptional control signals of a eukaryotic protein-coding gene. *Science* 217: 316-324.

Melton D. A., Kcieg P. A., Rebagliati M. R., Maniatis T., Zion K., and Green M. R. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* 12: 7035-7056.

Ng P., Wei C. L., Sung W. K., Chin K. P., Lipovich L., Ang C. C., Gupta S., Shahab A., Ridwan A., Wodg C. H., et al. 2005. Gene identification signature (GIS) analysis for transcriptome characterizalion and genome annotation. *Nat. Methods* 2: 105-111.

Pesole G., Liuni S., Grillo G., Licciulli F., Mignone F., Gissi C., and Saccone C. 2002. UTRdb and UTRsite: Specialized databases of sequences and functional elements of 5′ and 3′ untranslated regions of eukaryotic mRNAs. *Nucleic Acids Res.* 30: 335-340.

Sambrook J. and Russell D. W. 2001. *Molecular cloning: A laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schaefer B. C. 1995. Revolutions in rapid amplification of cDNA ends: New strategies for polymerase chain reaction cloning of full- length cDNA ends. *Anal. Biochem.* 227: 255-273.

Scherf M., Klingenhoff A., and Werner T. 2000. Highly specific localization of promoter regions in large genomic sequences by Promotertnspector: A novel context analysis approach. *J. Mol. Biol.* 297: 599-606.

Shiraki T., Kondo S., Katayama S., Waki K., Kasukawa T., Kawaji H., Kodzius R., Watahiki A., Nakamura M., Arakawa T., et al. 2003. Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. *Proc. Natl. Acad. Sci.* 100: 15776-15781.

Slansky J. E. and Farnham P. J. 1996. Transcriptional regulation of the dihydrofolate reductase gene. *BioEssays* 18: 55-62.

Smale S. T. and Baltimore D. 1989. The "initiator" as a transcription control element. *Cell* 57: 103-113.

Smale S. T. and Kadonaga J. T. 2003. The RNA polymerase II core promoter. *Annu. Rev. Biochem.* 72: 449-479.

Stapleton M., Liao G., Brokstein P., Hong L., Carninci P., Shiraki T., Hayashizaki Y., Champe M., Pacleb J., Wan K., et al. 2002. The Drosophila gene collection: Identification of putative full-length cDNAs for 70% of D. melanogaster genes. *Genome Res.* 12: 1294-1300.

Sugahara Y., Carninci P., Itoh M., Shibata K., Konno H., Endo T., Muramatsu M., and Hayashizaki Y. 2001. Comparative evaluation of 5-end-sequence quality of clones in CAP trapper and other full-length-cDNA libraries. *Gene* 263: 93-102.

Takai D. and Jones P. A. 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl. Acad. Sci.* 99: 3740-3745.

Wakaguri H., Yamashita R., Suzuki Y., Sugano S., and Nakai K. 2008. DBTSS: Database of transcription start sites, progress report 2008. *Nucleic Acids Res.* 36: D97-D 101.

Zhang M. Q. 1998. Identification of human gene core promoters in silico. *Genome Res.* 8: 319-326.

Zhang M. Q. 2007. Computational analyses of eukaryotic promoters. *BMC Bioinformatics* (suppl. 6) 8: S3.

Zhao X., Xuan Z., and Zhang M. Q. 2007. Boosting with stumps for predicting transcription stact sites. *Genome Bi-*

ol. 8: R17.
Zinn K., DiMaio D., and Maniatis T. 1983. Identification of two distinct regulatory regions adjacent to the human β interferon gene. *Cell* 34: 865-879.

网络资源

http://bioinfo2. ugr. es/CpGcluster/: CpGcluster 网站使用基于距离的算法检测 CpG 岛。

http://bioweb. pasteur. fr/seqanal/interfaces/cpgplot. html: CpGPlot 提供输入序列的 CpG 富集区的图形输出。

http://dbtss. hgc. jpl: DataBase 转录起始位点数据库汇集了实验确定的全长 cDNA 的 5′ 端序列，以及解释这些序列的分析工具。

http://www. dnastar. com/products/primerselect. php: 可商业化获得的 PrimerSelect 寡核苷酸设计信息，以及可从 DNASTAR 公司获得的分析程序。

http://frodo. wi. mit. edu. /cgi-bin/primer3/primer3www. cgi: Primer3 寡核苷酸设计程序的直接链接。

http://frontend. bioinfo. rpi. edu/zukerm/home. html: Zuker 实验室主页，提供到"mfold"算法及其他二级结构预测程序的链接。

http://genome. gsc. riken. jp/: RIKEN 基因组探索研究组（RIKEN Genome Exploration Research Group）的主页。

http://125. itba. mi. cnr. it/cgi-bin/wwwcpg. pl: WebGene 站点的 CpG 岛分析功能的直接链接。

http://www. oligo. net: 可商业化获得的由 Molecular Biology Insights 提供的寡聚引物分析软件（Oligo Primer Analysis Software）信息。

http://www. uscnorris. com/cpgislands2/cpg. aspx: CpG 岛搜索者（CpG Island Searcher）筛选提交 DNA 序列中符合使用者选择标准的 CpG 岛。

4　启动子分析的功能性分析方法

要点

- 功能性分析方法的发展是启动子分析中的重要早期步骤。
- 瞬时转染和稳定转染分析、体外转录分析、转基因和同源重组分析各有其不同的优缺点。
- 染色质免疫沉淀（ChIP）和 ChIP-chip 分析可能提供一些关于调控一个或一组启动子的机制的信息，但需进行功能性研究来确认 ChIP 的结果。

引言和概述

当开始基因转录调控机制的分析时，一个关键的目标是发展一种分析方法，并将其用于测量相关顺式作用控制区的活性（见第 2 章）。有了适当的功能性分析方法，才可能评定一个控制区精确模拟内源性基因表达模式的能力。然后，可将突变引入控制区，

以鉴定重要的调控元件，并最终鉴定重要的转录因子。在缺乏功能性分析方法时，非常难于（不是不可能）鉴定和评定DNA序列元件的功能相关性，并可能有助于基因调控的蛋白质。随着染色质免疫沉淀（ChIP）分析和ChIP-chip分析的出现，一个有效的功能性分析方法的重要性变得尤其明显。在ChIP-chip分析中，叠连阵列用于在全基因组基础上鉴定控制区中因子的存在。这些研究中的许多研究都假定结合与功能相关。然而，正如第8章中所详尽描述的，有必要用功能性实验验证ChIP的结果，以及这些功能性实验检验目的因子结合位点的功能相关性及其对基因调控的相对贡献。

有数种功能性分析方法已用于研究转录调控，最常用的方法是瞬时转染分析，在这种分析中，通过几种转染程序之一将含有目的控制区的质粒导入培养细胞（图4.1）。

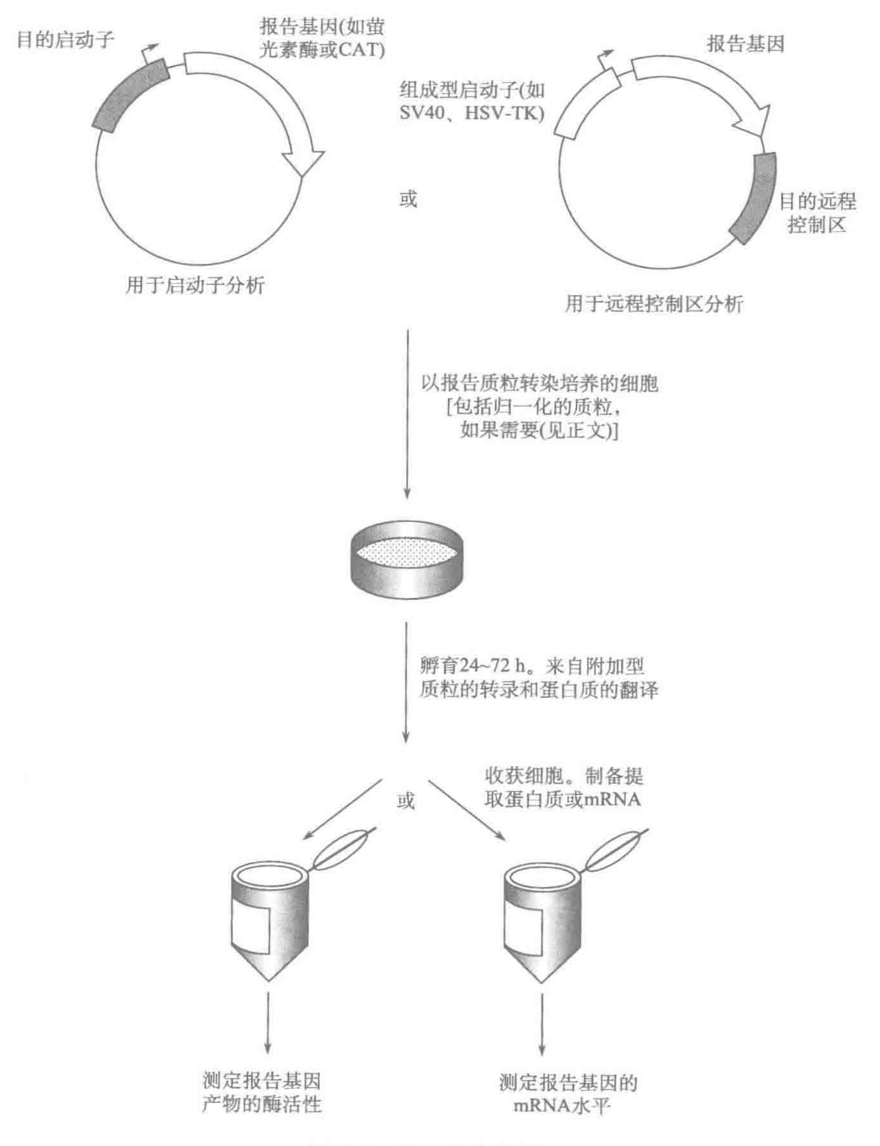


图4.1 瞬时转染分析。

一般情况下，控制区调控所谓“报告基因”的转录。报告基因是一个已知基因，它的mRNA和蛋白质水平可轻而易举地精确测量。如果目的调控区是一个启动子，它被置于报告基因的最近上游，以便该启动子驱动报告基因转录。如果目的控制区是在启动子远处起作用的增强子或其他控制区，一个良好表征的启动子通常被置于报告基因的上游，同时将增强子插入到启动子上游或报告基因下游。一般首选下游位置，因为将增强子放置在靠近启动子的区域可改变其特征，如个别增强子元件可以异常的方式与附近启动子元件协作。

在用产生的质粒转染培养细胞后的特定时间点，通过测定从报告基因上合成的mRNA或蛋白质，评价控制区的活性。这种分析被认为是瞬时的，因为质粒仍然以附加体的形式存在，并很少整合进宿主基因组。因此，mRNA或蛋白质产物必须在短期内进行测定，时间为1～3天不等。否则，随着细胞的生长和分裂，质粒会被降解或稀释。虽然瞬时转染分析法具有多种局限性（见下文），但该方法通常用于对调控基因表达的顺式作用DNA序列和反式因子进行初步分析。

对于某些控制区，另类分析是绝对必要的，或另类分析比瞬时分析更可取。例如，对于在瞬时分析中未表现出预期活性的控制区，或依赖于特定染色质结构的控制区（见第1章和下文）。稳定转染分析就是一种另类分析，在该分析中含有报告基因和目的控制区的质粒被稳定地整合进基因组（图4.2）。还需要由组成型启动子控制的抗药性基因（如显性选择标记基因），该基因可以和报告基因在同一个质粒上，也可以在不同的质粒上。

将含有报告基因和抗药性基因的质粒转染进培养细胞。通过向生长培养基中补加能杀死不稳定表达抗药性基因的细胞的药物，筛选质粒被稳定整合进染色体的细胞。在多数情况下，细胞内数个质粒分子彼此相连，并且每个多聚体整合在基因组中完全随机的位置。因此，每个稳定转染的细胞都具有一个或多个独特的整合位点，同时每个整合位点含有数个连接在一起的质粒分子。然后，通过测定稳定转染细胞池中报告基因的活性，确定目的控制区的活性。另外，可以分离并扩增数个单个的细胞克隆，以便每个克隆培养物中的所有细胞具有整合进一个或多个相同染色体位置的相同数量的质粒分子。

稳定转染分析的一个变通方法是使用附加型报告质粒，该质粒通过复制的细胞周期调控被稳定维持在恒定的拷贝数。为了维持附加体的存在，报告质粒中必须包括来自适当病毒的复制起点，以及由组成型活跃启动子控制的抗药性基因。转染后，将药物加至培养基中以选择被转染的细胞，并测定报告基因活性以确定目的控制区的活性。

除了转染方法外，还可以考虑使用三种其他的功能分析方法。一种是体外转录分析，在这种分析中，在无细胞的提取物中测定控制区的功能（图4.3）。另一种用于研究转录调控的分析方法是转基因分析，在这种分析中，将报告基因和目的控制区稳定地整合进一种动物基因组中完全随机的位置，以测量控制区在更自然的环境中的精确活性（图4.4）。用于划定功能控制区和控制元件的最后一种分析方法是同源重组分析，在这种分析中，在培养细胞或动物的基因组中对内源基因进行操作（图4.5）。

下面的部分描述在选择适当的功能性分析方法时必须考虑的问题，然后，详细讨论瞬时和稳定转染分析的策略问题及方法。体外转录分析将在第12章中进行单独讨论，转基因分析将在第5章进行更充分的讨论。同源重组分析的具体策略和技术超出了本书

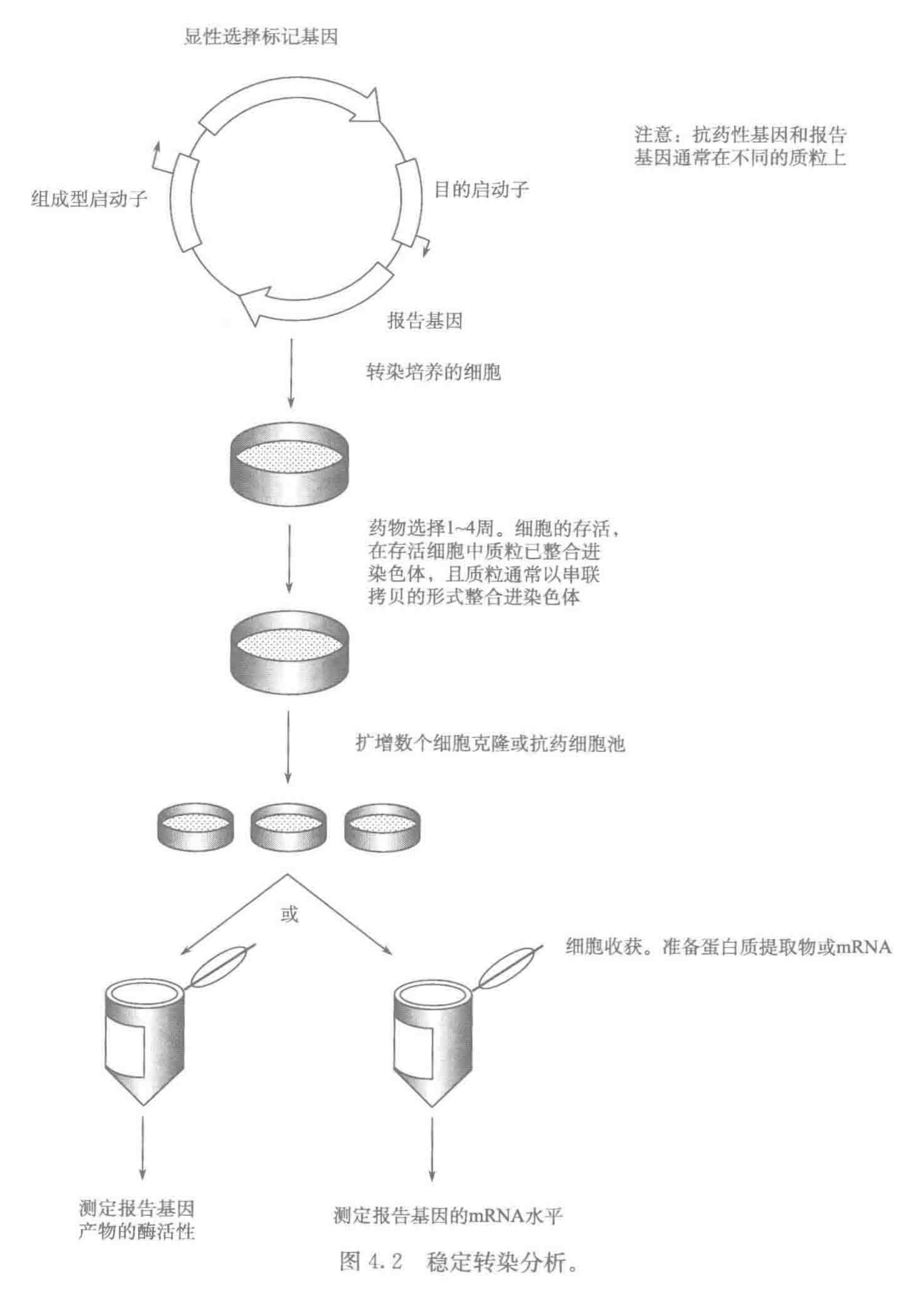


图 4.2 稳定转染分析。

范围，不予讨论。

实验策略

选择分析方法：各种分析方法的优缺点

现有的用于分析控制区的功能性分析方法的质量，常常决定该分析成功的水平，因

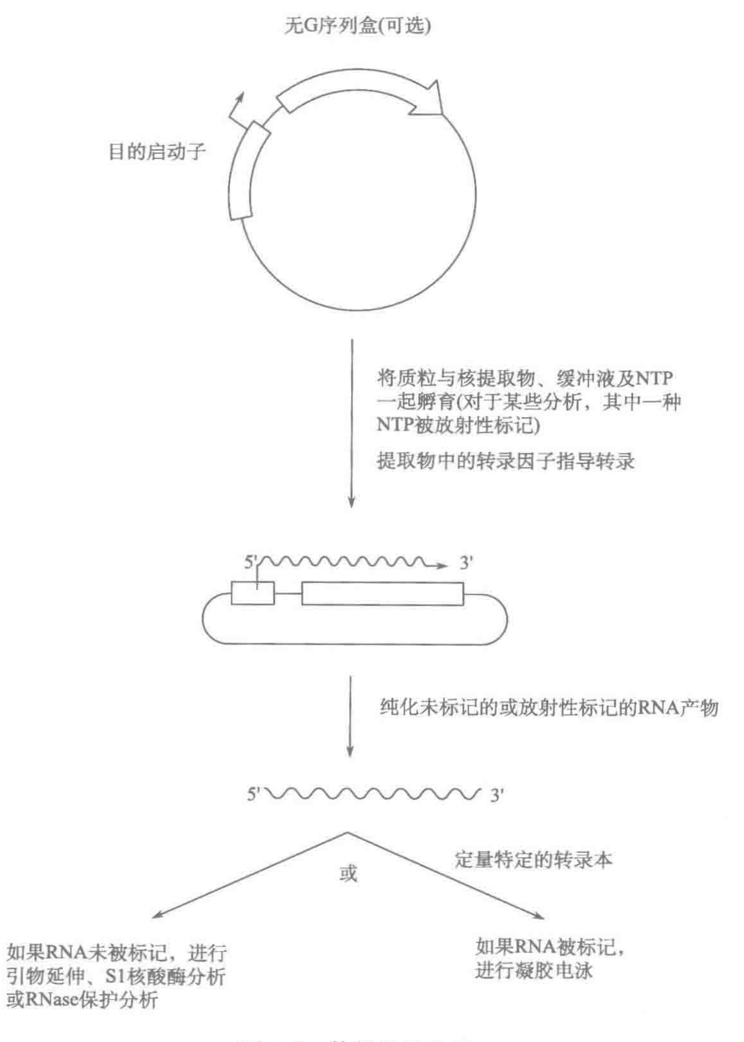


图 4.3　体外转录分析。

此，在选择分析方法时必须深思熟虑。遗憾的是，没有一种单一的分析方法是理想的，并且每种分析方法都有其重大的局限性。在任何情况下，一旦获得使人信服的证据，表明选择的功能性分析方法模仿或至少部分地模仿目的启动子的正确调控，那么在以分析启动子为目标的研究中，下一步是鉴定对启动子功能重要的顺式作用 DNA 序列。第 6 章概述了逼近这一目标的策略。

瞬时转染分析

对于大多数基因来说，瞬时转染显现了许多优势，从而使其成为首选的分析方法(至少在分析的初始阶段是这样)。瞬时转染分析操作快速简单，结果易于定量。当考虑到其他分析方法的局限性时，这些优势的重要性变得显而易见（见下文）。然而，瞬时分析有两个主要局限性。首先，在转染细胞中，质粒以一种人造构象和拷贝数存在，这

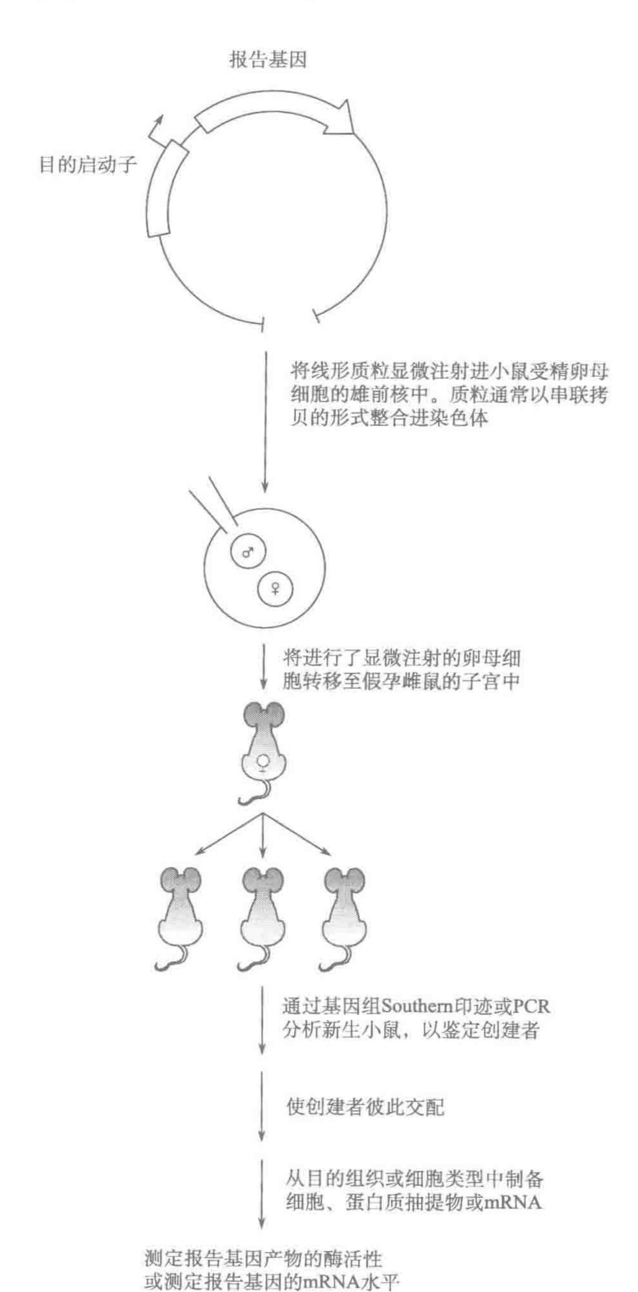


图 4.4 转基因分析（经许可，改编自 Janeway and Travers 1996，© Garland Science/Taylor & Francis，LLC）。

种构象和拷贝数可能导致特定控制元件的失活或功能异常（Mercola et al. 1985；Smith and Hager 1997），然而，不能过分强调这种局限性。数百或数千的质粒分子可进入转染细胞；因为重要转录因子以有限的数量存在，所以只有极少数的质粒可能与控制区行

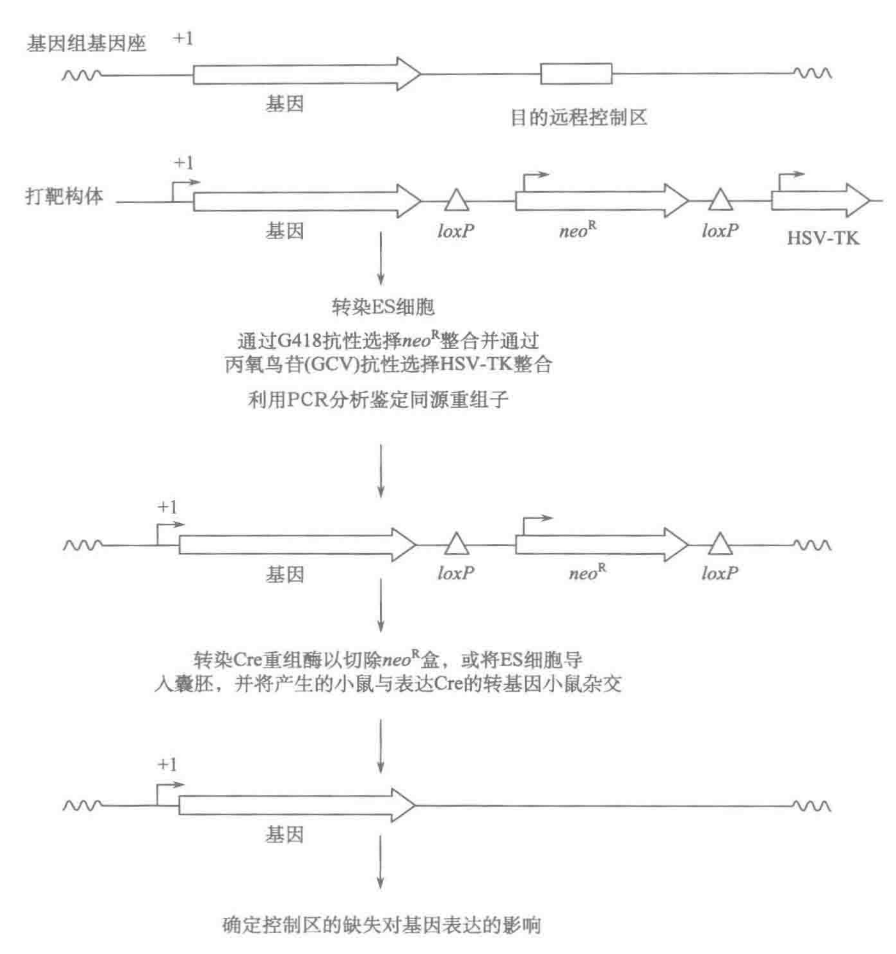


图 4.5　同源重组分析。

使正常功能所需的全部互为补充的蛋白质相结合。事实上，大量的转染报告质粒能抑制特定控制元件和 DNA 结合蛋白的活性（见 Mercola et al. 1985）。

质粒的附加性和不可复制性是引起控制区或元件功能异常的另一个人为特征。在这种情况下，异常的功能显然产生于一种事实，即质粒不在合适的染色质构象中（Smith and Hager 1997）。基因座控制区（LCR）和沉默子是两种通常不能在瞬时分析中起作用的控制区（见第 1 章、第 5 章和下文）。有些启动子和增强子还依赖于特定的染色质结构，并且因此在瞬时分析中不能被正确调控。即使在瞬时分析中启动子或增强子看似被正确调控，但很可能只是控制区中单个 DNA 元件的一个子集在该分析中正在发挥作用。通过研究对检测到的活性负责的 DNA 元件，可以获得很多知识，但是有关调控机制的完整的知识最终需要使用更先进的分析。

瞬时分析的另一个局限性是它不能用于测量一种特殊区域的活性，这种区域需要一段诱导或分化时间，而这段时间又超过了瞬时分析的时间限制（48～72 h）。例如，鼠红白血病细胞可被诱导经历一个通常需要数天时间的红细胞分化过程。因此，对于监测

这种细胞分化期间和分化后控制区的活性来说，瞬时转染分析是没有用的（Wright et al. 1953，Forrester et al. 1989）。

通过整合进宿主染色体的稳定转染分析

最常见的稳定转染分析形式是将报告质粒整合进宿主染色体。与通过瞬时转染分析相比，该分析方法的主要优势是正在分析的控制区/报告基因通常处于更自然的染色质构型，具有更自然的拷贝数（见 Smith and Hager 1997 及其中的参考文献）。由于这些特征使控制区更准确地模拟其正常功能，所以，当在瞬时分析中研究无活性或不显示其预期活性的控制区时，稳定分析会很有用。此外，对于对转染比较有抗性的细胞类型，整合了转染 DNA 的细胞的选择提供了另一优势：尽管吸收 DNA 的细胞的数量可能太少而不能通过瞬时转染分析，但是，在稳定转染分析中，转染的细胞可被选择性扩增，从而导致增强的转录信号。

鼠末端转移酶启动子为这些优势提供了范例。即使存在异源增强子，并且即使当使用最敏感的报告基因分析方法时，该启动子在标准的瞬时转染分析中也不能显著地影响报告基因活性（Lo et al. 1991）。相比之下，来自稳定转染的启动子/报告基因质粒的转录本易于检测（Trinh et al. 2001）。据推测，瞬时分析失败的原因是必须使用的细胞系转染效率差和（或）附加型质粒的拷贝数高。F9 胚胎癌细胞中的 c-*jun* 启动子提供了另一范例（Smith and Hager 1997）。瞬时转染后该启动子高度活跃，但只有在稳定整合后才观察到正确的调控。当分析只有整合进染色体时才能起作用的控制区（如 LCR 和沉默子）时，稳定分析才会特别有用（见第 5 章）。调控 β-珠蛋白基因座的 LCR 提供了一个例子，因为 LCR 及其多数亚结构域在稳定转染分析和转基因小鼠分析中是有活性的，但在瞬时转染分析中无活性（Blom van Assendelft et al. 1989；Talbot and Grosveld 1991）。

稳定转染分析的最后一个优势是对于随后的转录分析没有时间限制。因此，对于研究需要相对较长孵育期的分化和诱导事件，稳定转染分析是有用的（Wright et al. 1983；Forrester et al. 1989；Trinh et al. 2001）。在通过多种外部试剂测量报告基因活性的诱导或抑制时，报告基因的稳定整合也是有利的。关于瞬时分析，需要在每个实验之前转染细胞。关于稳定分析，可向已经建立的细胞系加入诱导物或抑制物。

然而，涉及染色体整合的稳定转染分析比瞬时转染分析更难且更费时。因此，稳定转染分析通常需要数周，而不是完成瞬时转染分析所需的 48～60 h，所需的大多数时间用于药物筛选和细胞扩增。稳定转染分析的另一个局限是，报告基因的转录活性和调控受到其整合位点的固有可及性的强烈影响。术语“可及性”定义不清，但是可及性可能是指邻近活性基因和非活性基因的染色质结构的差异。据认为，只有报告基因被整合进邻近活跃转录的内源基因的可及的染色质结构中，整合的报告基因的活性才将很强。相反，抑制性染色质结构附近的整合可导致整合质粒的沉默。某些控制区（如 LCR 和绝缘子）能够克服整合位点可及性的差异。此外，将绝缘子或边界元件置于报告质粒两侧，使整合位点的效应最小化（见第 1 章和第 5 章及下文）。尽管如此，由于这些整合位点的变化，通常有必要表征来源于每个稳定转染实验的多个克隆，并慎重解释获得的结果。例如，稳定转染分析通常用于分析启动子和增强子突变体，以划定生理相关性控

制元件。遗憾的是，实际上不可能分析活性只改变 2～3 倍的突变体，因为克隆间的变异将超过此数量。因此，对于确定严重改变控制区活性的突变体的效应，稳定转染分析可能是有用的，但对于确定具有相对较小效应的突变体的效应，稳定转染分析的作用不大。

维持附加型质粒的稳定转染

一种替代性稳定转染分析中使用的是含有来源于受细胞周期和拷贝数调控的病毒复制起点的报告质粒，如 Epstein-Barr 病毒（EBV）和小牛乳头瘤病毒（Kushner et al. 1982；Zinn et al. 1982；Yates et al. 1985；Ohe et al. 1995）。该分析方法具有许多与更常见的基因组整合型稳定转染分析相同的优缺点，但它还具有其自身的一套优势和缺点。其主要优势在于基因转录不受染色体整合位点引起的变异的影响，因此，克隆之间得到的结果应更有可比性。而且，虽然转染的质粒仍然是附加型的，但其复制能力使其像目的内源控制区一样组装成染色质结构（Beato 1996；Smith and Hager 1997 及其中的参考文献）。

该分析方法的重要局限性是，维持附加型质粒可因细胞繁殖而不稳定，并经历重组事件（Chittenden et al. 1989）。如果通过重组产生的新的附加型种类赋予细胞生长优势，那么该新种最终将会在培养物中成为优势附加型质粒。或者，如果重排使复制起点失活，那么产生的质粒会整合进基因组，并从其基因组位置赋予抗药性。由于这些局限性，必须频繁而广泛地在分子水平上表征通过附加型质粒产生的稳定细胞系，以确保质粒仍然是附加型的，且未发生重排。鉴于此，对于染色质结构对基因活化影响的专门研究，主要推荐附加型稳定转染分析。使用该分析方法的突出实例是核小体调控小鼠乳腺瘤病毒启动子的分析（Bartsch et al. 1996，Beato 1996）。一个更新的例子涉及由 ATP 依赖的核小体重塑复合物激活的干扰素靶基因的表征（Liu et al. 2002）。本书将不对该分析方法作进一步讨论，相关方法和策略的更多信息可参考 Beato 和 Zhao 实验室发表的论文。

体外转录分析

对于特别抗转染的细胞类型、在转染分析中不能产生可检测活性的启动子，或基因调控的高级研究，体外转录分析是有用的。该分析方法的显著优势是不需要高效的转染条件，此外，它可以向体外反应体系加入抗体，而且可以加入或去除转录因子，以评价特定转录因子的功能（见第 12 章）。这里所描述的各种分析方法中，体外分析最迅速，一天之内就可以完成整个实验。

体外分析的主要局限是参与转录调控的因子很有可能无活性、活性弱或功能异常。许多转录因子在提取物制备期间不能保持其完全的活性，或者完全不能在体外刺激转录。一种可能的解释是因子在核提取物中的浓度小于其在细胞核中的浓度。而且，某些因子依赖于核小体的存在，或者需要其他染色质关联蛋白。核小体也可以通过抑制基础转录增强对因子的应答性。与使用裸模板的研究相比，使用组装成核小体的模板获得调控型体外转录的研究已经取得更大的成功（Barton et al. 1993，Pazin et al. 1996，Kraus and Kadonaga 1998；Agalioti et al. 2000；Lomvardas and Thanos 2002）。尽管有这些局限性，体外分析已用于再现几种启动子的受调控活性或组织特异性活性，这些

启动子包括免疫球蛋白（Ig）启动子（Pierani et al. 1990）、人免疫缺陷病毒类型 1 启动子（Pazin et al. 1996）、白细胞介素-2 启动子（Flanagan and Grabtree 1992）、末端转移酶启动子（Garraway et al. 1996）和β-干扰素启动子（Agalioti et al. 2000；Lomvardas and Thanos 2002）。

与使用启动子获得的成功相比，远程控制区在体外很少发挥作用（Carey et al. 1990）。为检测远程增强子的活性，通常有必要将其插入紧接启动子的上游，这样会使某些增强子元件与附近的启动子元件发生不正确的协同作用，通过这种作用，导致异常功能。对远程增强子的进一步体外研究将最有可能涉及核小体重建模板，这些模板似乎从远距离促进增强子发挥其功能。然而，该策略仅在少数研究中被采用（Kadonaga 1991；Laybourn and Kadonaga 1992；Barton et al. 1993，1997），而且其对增强子功能分析一般有用的程度仍然有待确定。鉴于上述考虑，建议将体外分析用于补充体内分析，或者在体内表征了特定控制元件和蛋白质的功能和特性之后，再通过体外分析研究其特性。只有在发展体内功能性分析的尝试不成功时，才推荐将体外分析作为初步或主要的分析方法。

转基因分析

由于本书主要论述哺乳动物系统，因此将转基因小鼠作为典型动物分析加以讨论。转基因小鼠已经多年用于确证在转染分析中鉴定的特定控制区或元件在动物体内确实很重要。Ig μ 增强子的转基因研究提供了这样一个例子（Ernst and Smale 1995）。Ig μ 增强子是用瞬时转染和稳定转染分析发现的，并预测其可使 Ig μ 转录具有 B 细胞特异性（Banerji et al. 1983；Gilles et al. 1983）。然而，为了确定该增强子的准确组织特异性，需要进行后续转基因研究（Jenuwein and Grosschedl 1991）。也可以用在单个 Ig μ 控制元件中含有突变的构体产生转基因小鼠，其中这些构体最初是在瞬时转染实验中鉴定的（Jenuwein and Grosschedl 1991；Annweiler et al. 1992），这些转基因实验确证了控制元件的重要性。

转基因分析也已被用于鉴定在瞬时转染或稳定转染分析中通常不起作用的 LCR。LCR 刺激高水平的、不依赖于整合位点的转录（见第 1 章和第 5 章）。虽然稳定转染分析能提供整合位点不依赖性的初步证据，但是获得的结果不太确切，因为需要共整合的抗药性基因。第 5 章更详细地讨论了研究 LCR 的分析方法。

转基因分析的主要优点是它能洞察在动物体背景中控制区或单个 DNA 元件的功能。此外，与稳定转染分析不同，转基因分析不需要在单独的启动子和增强子控制下抗药性基因的共整合。这是有利的，因为这样的第二个转录序列盒的存在经常影响稳定转染实验的结果。

转基因分析的缺点是转基因整合在基因组内完全随机的位置。与稳定转染分析类似，周围序列和染色质环境可强烈影响所研究的控制区的活性。目前已发展了指导转基因整合进确定的基因组位置的策略（如见 Beard et al. 2006），这些策略有可能消除由整合位点的变化造成的问题。尽管如此，除非使用这些策略之一，否则将有必要常规分析多个独立的转基因系，以确定整合位点对结果影响的程度。

在使用转基因小鼠进行的转录调控分析中，一个主要的进展是已经发展了用于操作

细菌人工染色体（BAC）的技术，以便像转基因一样有效地研究细菌人工染色体。这些构体含有通过大肠杆菌中的同源重组被插入目的基因座的报告基因。通过使用彻底越过目的基因边界的大 BAC，通常会使整合位点的变化最小化，并且使转基因有相对高的概率重现在内源基因座上观察到的调控事件。由于 BAC 转基因对于定位远程控制区特别有用，所以将在第 5 章对其进行更为深入的讨论。

同源重组分析

在细胞系或动物中，通过同源重组对内源控制区或控制元件进行突变（Capecchi 1989），能够提供该区域或元件对基因调控是重要的最确凿的证据。同源重组很少用作转录调控研究的功能分析方法，因为该方法相对缓慢和困难。其困难来自于需要操作 DNA 大片段以产生重组底物，同时来自于挑选通过同源重组整合了底物的稀有细胞克隆的复杂性。当在小鼠而不是细胞系中进行同源重组时，困难来自两个方面，一是需要维持健康的胚胎干细胞（ES），二是需要获得重组等位基因在种系传递。由于这些局限性，直到最近同源重组分析也仅用于少数转录调控研究，主要用于确定特定增强子区对内源基因表达的重要性。两个早期的例子涉及 T 细胞受体 β 增强子的缺失（Bones et al. 1996）和 Ig κ 内含型增强子的缺失（Xu et al. 1996）。

尽管事实上同源重组分析仍然昂贵且困难（相对于上文所描述的分析方法），但该方法目前更为频繁地被使用，因为主要的技术进步使其比以前费时更少。特别是，全基因组序列信息的获得已经在很大程度上消除了在产生打靶载体前对基因座广泛的限制性作图的需要。而且，由于高保真长程 PCR（聚合酶链反应）的发展和全基因组序列的获得，现在可通过 PCR 而不是通过克隆限制性大片段制备打靶载体。最后，用于 ES 细胞转染和维持以及囊胚注射的高品质核心设备的不断获得，使得该方法对于一般分子生物学实验室比以前更加可行。

由于这些进步，已经使用同源重组分析了许多鼠源控制区的功能。对于鼠源 *Il4* 基因，通过打靶删除已分析了至少 4 个不同的远程控制区，其他控制区的打靶载体正在构建中（Mohrs et al. 2001；Solymar et al. 2002；Ansel et al. 2004；Lee et al. 2005）。打靶载体通常在将被删除的控制区两侧含有 3 kb 或更长的同源臂，同时从载体里排除将被删除的控制区。同源臂通常位于新霉素（neo）抗性表达盒的两侧，该表达盒用于挑选已整合打靶构体的 ES 细胞克隆。用 Southern 印迹和（或）PCR 鉴定抗药性 ES 细胞克隆，相对于基因组内其他位置的随机整合，这些克隆中已通过同源重组在目的基因座中整合了打靶构体。由于用于驱动新霉素抗性基因表达的启动子和增强子可能影响目的内源基因的转录，所以将 *loxP* 位点置于新霉素盒的两侧，从而通过 Cre 介导的重组切割新霉素盒（Sternberg and Hamilton 1981；Gu et al. 1993）。尽管也可以在与 Cre 表达系杂交产生突变小鼠后进行 Cre 重组，但是通常通过将 Cre 表达质粒转染到已成功同源重组的 ES 细胞中实现切割。

尽管事实上同源重组分析将会提供非常相关的信息，但它很少用于分析单个控制元件的重要性。而且，其主要局限性是所需的工作量。尽管如此，同源重组已用于研究 *Cd4* 沉默子中几个单个元件的重要性（Taniuchi et al. 2002）。在这种情况下，将预期的突变整合进新霉素盒两侧的其中一个同源臂中；将抗药盒整合到沉默子附近，并且之

后通过 Cre 介导的重组去除抗药盒。

如果目的是将几个不同的突变分别导入特定控制区，那么另一种方法是使用双替换策略（Wu et al. 1994；van der Weyden et al. 2002）。该方法中，第一步，将单纯疱疹病毒胸苷激酶基因（HSV-TK）连同新霉素抗性基因一起导入目的控制区。第二步，使用含有完整控制区和侧翼序列的第二打靶载体进行同源重组，其中控制区只有最终期望的置换突变或微小缺失突变。通过同源重组整合打靶载体将导致 HSV-TK 和新霉素盒的切除；可在含有丙氧鸟苷的培养基中选择成功实现这一重组事件的细胞，丙氧鸟苷杀死所有继续表达 HSV-TK 基因的细胞。因此，用含有整合 HSV-TK 和新霉素盒的 ES 细胞系作为出发点，可将许多不同的突变导入目的控制区，无需随后的新霉素盒切除。

将来，当用于在哺乳动物细胞中进行同源重组的方法变得更容易且耗时更少时，同源重组作为评估单个 DNA 序列元件功能的工具可能会变得越来越有价值和常见。事实上，同源重组有可能进一步提高下述过程的速度：通过在体外将含有靶向突变的 ES 细胞分化为期望得到的细胞类型，可产生并分析内源控制区中的一系列突变。使用该策略，应该有可能确定突变对组织特异性基因的转录活性的影响，无需以每种突变 ES 细胞系产生小鼠。

瞬时转染分析

为发展瞬时转染分析以分析基因的启动子，必须完成下列步骤。

1. 确定培养细胞的来源。
2. 确立将 DNA 导入细胞的转染程序。
3. 选择报告基因和相应的报告分析方法。
4. 鉴定目的基因的假定启动子区（见第 3 章）。
5. 将启动子插入到适当载体中的报告基因上游。

完成这些初始步骤后，就可以尝试瞬时转染实验，以确定启动子依赖的报告基因活性是否是可检测的。

细胞

用于瞬时转染分析的细胞应符合两项标准：①该细胞应表达衍生出目的启动子的内源基因；②在完成实验所需的 2～3 天内，在培养物中应能保持恒定数量的细胞。此外，必须有将质粒 DNA 转染进细胞的程序。转化的（无限增殖的）细胞系通常是瞬时转染分析的最佳选择，因为来源于许多细胞类型的细胞系易于从商业来源和学术来源获得，其中，商业来源包括美国模式培养物保藏所（American Type Culture Collection）（http://www.atcc.org/）、由科里尔医学研究所（Coriell Institute for Medical Research）运营的国家普通医学研究所（National Institute of General Medical Sciences）细胞库（http://ccr.coriell.org/Sections/Collections/NIGMS/?SsId=8）等。另外，转化细胞系相对易于维持，而且，对于许多细胞系来说，已经建立并发表了有效的转染程序。相比之下，在培养基中生长的原代细胞群体通常难以分离和维持，且对于常用的转染程序有抗性。如果不能获得表达内源目的基因的转化细胞系，就可能有必要确定使用原代细胞群的可行性，或者通过癌基因转化建立无限增殖的细胞系。

转染程序

如果得到了表达目的内源基因的细胞系但还未对其确立转染条件，就必须先确定最佳的转染条件。数种转染方法已用于瞬时启动子分析，包括磷酸钙转染、DEAE-葡聚糖转染、电穿孔转染、脂质体转染、使用其他阳离子试剂的转染及原生质体融合的转染。专题 4.1 包含了对最常用的程序的介绍。有关转染程序的更详细信息可从引用的具体文献获得，也可从 Ausubel 等（1994，pp. 9.1-9.5）、Spector 等（1988）和 Sambrook and Russell（2001，pp. 16.1-16.62）等文献中获得。

为了确定哪种转染程序对于给定的细胞系最可能成功，可进行文献检索以确定哪种方法已经成功地用于该细胞系或相关细胞系（即来源于相似组织的细胞系）。例如，已通过 DEAE-葡聚糖、电穿孔或脂质体等方法成功转染了多数淋巴细胞系，但是很少用到磷酸钙。因此，当确立新淋巴细胞系的转染条件时，磷酸钙将是糟糕的选择。同样重要的是要注意应根据目的基因选择转染策略。例如，使用磷酸钙转染策略可能难以研究钙诱导型基因。

专题 4.1　常用的转染方法

磷酸钙法

将 DNA 和氯化钙在磷酸缓冲液中混合，导致 DNA-磷酸钙沉淀的形成；之后，将 DNA-磷酸钙沉淀铺在细胞上（Graham and van der Eb 1973），细胞通过吞噬作用将沉淀吸收。该程序广泛用于成纤维细胞和其他贴壁型细胞的瞬时转染和稳定转染实验。使用试剂级化学试剂容易进行该程序，或者，可以从商业来源购买为磷酸钙转染实验专门设计的试剂和试剂盒［如哺乳动物转染试剂盒-磷酸钙（Stratagene，200385）、磷酸钙转染试剂盒（Invitrogen，K2780-01）、ProFection 哺乳动物转染系统-磷酸钙（Promega，E1200）或 CalPhos 哺乳动物转染试剂盒（Clontech，631312）］。

DEAE-葡聚糖法

带有负电荷的 DNA 结合于阳离子型 DEAE-葡聚糖（McCutchan and Pagano 1968），将可溶性的复合物加到细胞上，细胞通过非特异性的吞噬作用将之吸收。该程序用于贴壁细胞和非贴壁细胞的瞬时转染。然而，该程序很少用于稳定转染实验，因为其毒性阻止转染细胞的长期维持。和磷酸钙法一样，使用试剂级化学试剂可容易地进行 DEAE-葡聚糖转染，尽管也有商业化试剂盒，如 ProFection 哺乳动物转染系统 -DEAE-葡聚糖（Promega，E1210）和哺乳动物转染试剂盒（Stratagene，200285）。

电穿孔法

将 DNA 和细胞混合于依经验确定的缓冲液中，然后使混合物经受高压电击，这会在细胞膜上诱导或稳定微孔，DNA 通过微孔进入细胞（Neumann et al. 1982；Ziemermann 1982）。该方法快速并可用于许多类型的贴壁细胞和非贴壁细胞。推荐使用 Bio-Rad Gene Pulser Xcell，因为它产生的指数波形和方波形增加可能导致目的细胞系有效转染的条件范围。Nucleofector 电穿孔仪（Amaxa Biosystems，AAD-1001）连同优化的缓冲液一起也可促进许多转化的细胞系和原代细胞群体的高效电穿孔。在任何情况

下，除了特定供应商提供或推荐的最佳质粒外，用自己的质粒测试转染效率很重要。

基于脂质体试剂的转染法

带有负电荷的 DNA 结合于阳离子型脂质化合物的正电荷末端（Feigner et al. 1987; Mannino and Gould-Fogerite 1988）。DNA 结合脂质和细胞膜相结合，使 DNA 进入细胞内部。对于某些细胞，脂质体产生特别高的转染效率；而对于其他细胞，转染效率和使用其他方法得到的转染效率相当或更差。几家不同的公司都出售专利脂质化合物，包括 CLONfectin（Clontech，631301）、Tfx-50 和 Transfectam（Promega，E1811 和 E1231）、LipoTAXI（Stratagene，204110）、DMRIE-C 和 Lipofectamine 2000（Invitrogen，10459-014 和 11668-019）、DOSPER 和 DOTAP（Roche Applied Science，11781995001 和 11811177001）和 Effectene Transfection Reagent（QIAGEN，301425）。这些化合物彼此之间稍有不同，因此，观察到的转染效率可能有非常大的差异。对于一个新的细胞系，难以预测哪一种化合物最有效。脂质体的缺点是相对于磷酸钙法和 DEAE-葡聚糖法花费更大。

多聚阳离子法

包括聚凝胺（Chaney et al. 1986）和聚乙烯亚胺（Boussif et al. 1995）在内的几种多聚阳离子可以促进质粒 DNA 向某些哺乳动物细胞系的有效转移。也可使用专利多聚阳离子，如 SuperFect 和 PolyFect 转染试剂（QIAGEN、301305 和 301105）。这些试剂易于使用，且会引起某些细胞系的有效转染。最有可能的是，这些化学药品的作用机制与 DEAE-葡聚糖类似。

当对新细胞系发展功能进行分析时，需要使用标准报告质粒测试和优化转染程序，标准报告质粒含有融合于报告基因的强病毒启动子和增强子，且报告基因有灵敏的分析方法（见下文）。在优化转染条件时，病毒启动子和增强子用于驱动报告基因的转录，因为它们比来自大多数细胞基因上的对照区域更强，并且因此将在优化尝试过程中赋予更高的敏感性。而且，病毒启动子和增强子含有异常大量的转录因子的结合位点，同时这些转录因子具有相当大的冗余性，因而它们在大多数细胞类型中都具有很强的活性（Schmidt et al. 1990）。Promega 的 pGL3-对照载体、pCAT3-对照载体和 pSV-β-半乳糖苷酶对照载体含有由 SV40 启动子和增强子控制的报告基因，对于优化分别使用蛍光素酶、氯霉素乙酰转移酶（CAT）、β-半乳糖苷酶分析法的转染条件很有用（见专题 4.2～专题 4.4）。诸如 Clontech 的 pAcGFP1-C1 质粒的绿色荧光蛋白（GFP）报告质粒也能用于优化转染条件（见专题 4.4）。当通过流式细胞仪监测 GFP 的荧光时，通常可获得有关转染细胞的百分比及转染群体中信号强度的详细信息。

用对照报告质粒确立了产生可检测活性的条件之后，应系统地变动各种转染参数，以优化报告基因的活性。对多数转染程序，最佳信号取决于转染效率和细胞死亡之间的平衡。因为严酷的条件通常使更多的 DNA 进入细胞，所以便于 DNA 吸收的理想条件通常使多数细胞死亡。相反，能引起 100%细胞成活率的条件对 DNA 吸收来说过于温和。使用促使某些细胞死亡而又保留足够健康细胞以驱动报告基因表达的条件，通常可以获得最佳的报告基因活性。在本章技术部分给出了一般被改变以优化转染条件的参数。

如果鉴定不到产生强报告活性的细胞系，可以尝试一些其他的瞬时转染策略。首先，如果得到的细胞系不易转染，则可用一种方法分离并浓缩少量已转染细胞，从而使这些细胞中报告基因的表达能够被监测（见专题 4.5）。其次，可以使用具有复制能力的质粒增加细胞中质粒的拷贝数，并因此增加报告基因的活性（见专题 4.6）。最后，根据目的调控区的性质，检测一种不表达相关内源基因的密切相关的细胞类型。相关的细胞类型中可能含有许多基因激活所需的转录因子。在相关细胞类型中，抑制内源基因的机制可能在人工瞬时分析中不发挥作用，因此启动子活性可能是可检测的。但是，该策略远非理想，一般应用作最后采用的办法。

作为将启动子-报告盒导入培养细胞的传统转染方法的替代方法，也已经发展了设计用于表达目的启动子控制下的报告基因的反转录病毒载体和腺病毒载体，并且这些载体的使用已取得了一些成功（Zheng et al. 2001；Zhu et al. 2001；Stirland et al. 2003；Leung et al. 2004）。以前使用病毒载体研究启动子和增强子活性的尝试通常不成功，因为未观察到控制区的正确调控。不过，最近已发展了提高成功率的策略。例如，有了反转录病毒载体，自我失活病毒的发展已经相当有利，因为自我失活破坏了 LTR 内的反转录控制区；在自我失活病毒中，反转录病毒的 5′ LTR（长末端重复）在染色体整合后被破坏（Leung et al. 2004）。

报告基因、载体和分析方法

除了确定用于瞬时转染的适当的细胞来源，也必须选择适当的报告基因，可以将报告基因的 mRNA、蛋白质或酶水平用作启动子活性的测量。在大多数情况下，使用报告基因监测启动子活性（而不用自然状态下受目的启动子控制的基因）的主要原因是，可用非常敏感和迅速的分析方法监测常用报告基因的表达。此外，报告基因的使用消除了辨别两种转录本的困难，即源于转染质粒的转录本和源于内源基因的转录本。

专题 4.2 萤光素酶报告基因分析

也许最常用的报告基因是来源于萤火虫 *Photinus pyralis* 的萤光素酶基因（de Wet et al. 1987）。该基因编码一个 61 kDa 的酶，它在 ATP、氧气和 Mg^{2+} 存在条件下氧化 D-萤光素，从而产生荧光产物，通过测定释放的光可定量测定该荧光产物（见 Promega 的《方案和应用指南》）。反应中加入辅酶 A 增强该分析的灵敏度并提供持续的光反应。萤光素酶分析极为迅速、简单、相对廉价、灵敏，并且具有宽的线性范围。Promega 声称萤光素酶分析比 CAT 分析敏感 30～1000 倍，而且提供大于 8 个数量级的线性应答。

Promega 提供多种萤光素酶报告载体。pGL2-基本载体和 pGL3-基本载体（Promega，E1641 和 E1751）在其多克隆位点下游含有萤光素酶基因，可将目的启动子插入多克隆位点中。为了减小归因于载体内的转录起始的萤光素酶背景活性，可将一个多腺苷酸化信号置于多克隆位点的上游。萤光素酶基因下游存在几个独特的限制性酶切位点，用于插入增强子或其他远程控制区（见第 5 章）。pGL2-增强子载体和 pGL3-增强子载体（Promega，E1621 和 E1771）在萤光素酶基因下游含有一个 SV40 增强子，以增强由弱启动子驱动的转录。pGL2-启动子载体和 pGL3-启动子载体（Promega，

E1631 和 E1761）在萤光素酶基因上游含有 SV40 启动子，这些载体可用于增强子区的鉴定和分析（见第 5 章）。pGL2-对照载体和 pGL3-对照载体（Promega，E1611 和 E1741）含有 SV40 启动子和增强子。pGL2 和 pGL3 的骨架不同，因为在 pGL3 的骨架上删除了已知转录因子的许多隐秘结合位点，目的是减少背景转录。Promega 目前提供另一家族的萤光素酶报告载体—— pGL4 系列，在该系列中删除了更多隐秘结合位点。pGL4 家族中的一些载体表达与 PEST 序列融合的萤光素酶蛋白，PEST 序列促进蛋白质降解，并且可进一步降低背景信号。

Promega 也提供了一套含有来源于海洋三色紫罗兰 *Renilla reniformis* 的萤光素酶基因的萤光素酶报告基因（Matthews et al. 1977；Liu et al. 1997），其所编码的酶有不同于 *Photinus* 酶的底物特异性，是在腔肠素而不是 D-萤光素存在的条件下发光。不同的特异性是 Promega 公司的双萤光素酶报告分析系统（dual-luciferase reporter assay system）的基础，该系统提供了比较目的启动子活性和对照启动子活性（目的是归一化）的快捷方法。例如，目的启动子可以插入 pGL3-基本载体，产生的质粒与 pRL-CMV 载体共转染细胞，pRL-CMV 载体含有巨细胞病毒（CMV）启动子/增强子控制下的 *Renilla* 萤光素酶基因。可以向转染细胞提取物中先添加 D-萤光素以测量 *Photinus* 萤光素酶活性，然后加入第二种试剂灭活 *Photinus* 酶并活化 *Renilla* 酶。

除了双萤光素酶报告分析系统外，也有用于 *Photinus* 萤光素酶活性传统测量的试剂盒［如萤光素分析系统（Promega，E1500）或高灵敏性萤光素酶报告基因分析（Roche Applied Science，11669893001 和 11814036001）］。这些试剂盒通常包括细胞裂解缓冲液、分析底物和分析缓冲液。通常在光度计中测量萤光素酶活性，但也可以在闪烁计数仪中测量其活性。

常用的报告基因。最常用的报告基因来源于昆虫［萤火虫萤光素酶（专题 4.2）］、原核生物［CAT（专题 4.3）和 β-半乳糖苷酶（专题 4.4）］或水母［GFP（专题 4.4）］，并且这些报告基因产生的酶活性或发出的荧光通常在大多数真核细胞中未被发现。目的启动子区被插入到报告基因编码序列的上游。因为报告载体通常含有 Kozak 一致序列（即包含真核翻译起始密码子的序列）和下游多聚腺苷酸化信号，由 RNA 聚合酶 II 合成的 mRNA 能有效地翻译成蛋白质。通常通过测量其酶活性或荧光强度可监测这些报告基因的表达，酶活性或荧光强度与启动子转录起始频率应大约成正比，因此检测报告基因的表达提供了相当准确的启动子活性评价。

专题 4.3 CAT 报告基因分析

第二种报告基因是 *E. coli* 转座子 CAT 基因（Gorman et al. 1982），该报告基因在以前常用，但是现在很少用，因为其敏感性远没有萤光素酶高且分析方法更繁琐。CAT 酶催化氯霉素的乙酸化，其乙酰基团由乙酰 CoA 或由商业化分析中的 *n*-丁酰 CoA 提供。可通过多种方法监测乙酰化的氯霉素，最常用的是将［^{14}C］氯霉素作为反应底

物，用薄层层析（TLC）法将乙酰化氯霉素从非乙酰化氯霉素分离出来后，通过放射自显影监测乙酰化作用（Gorman et al. 1982，图 4.6）。可通过 TLC 板的磷屏成像分析测量［^{14}C］氯霉素向乙酰化［^{14}C］氯霉素的转化率，该分析是将放射性斑点从 TLC 板上切下，在闪烁计数器中计数，或进行放射自显影图像的光密度分析。也可以不用 TLC，而是通过使用试剂级化学试剂进行有机提取和闪烁计数，以定量分析乙酰化^{14}C-标记的产物。或者，可从商业上获得含有多数试剂的 CAT 酶分析系统（Promega，E1000）。非放射性 ELISA（酶联免疫吸附分析）方法（Roche Applied Science，11363727001）也可用于检测 CAT 蛋白。Promega 提供与 pGL 萤光素酶载体相似的 pCAT3 系列载体（见专题 4.2），包括不含启动子或增强子的载体（pCAT3-Basic，E1871）、含有 SV40 启动子的载体（pCAT3-Promoter，E1861）、含有 SV40 增强子的载体（pCAT3-Enhancer，E1881）和既含有启动子又含有增强子的载体（pCAT3-Control，E1851）。

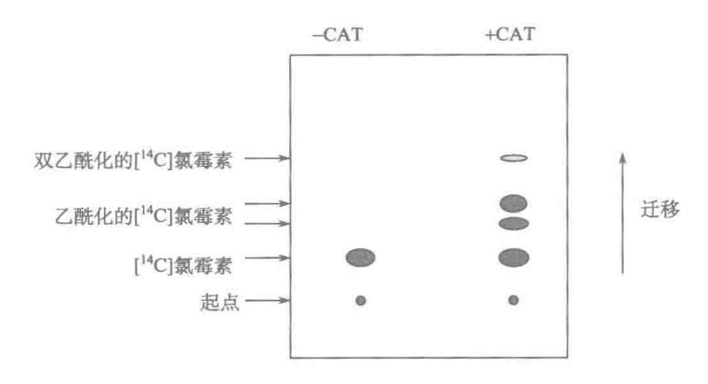


图 4.6　使用薄层层析的 CAT 分析。

专题 4.2～专题 4.4 描述了最常用的报告基因和分析方法，并列出了载体和分析试剂盒的来源。本章结尾给出了萤光素酶、CAT 和 β-半乳糖苷酶的分析方案，详细的信息和方案既可在引用的具体文献中查到，也可在 Ausubel 等（1994，单元 9.6）及 Sambrook 和 Russell（2001，第 16 章）的文献中查到。文献 Ausubel 等（1994）中包括有关大量报告基因分析方法详细信息的表格（单元 9.6.8）。此外，许多销售报告载体和分析方法的公司也提供相关信息。

*间接报告分析方法的优势。*上文提到的报告基因分析通常称为“间接”分析，因为将酶水平用作启动子活性的间接测量。相比之下，监测 mRNA 水平的分析（如引物延伸、S1 核酸酶或 RNase 保护）提供了启动子活性的更直接的测量（尽管不涉及 mRNA 的直接可视化）。间接分析的三个主要优势是简单、快速和灵敏。例如，间接分析法中最灵敏的萤光素酶分析比引物延伸法、RNase 保护分析法和 Northern 印迹分析灵敏约 1000 倍。萤光素酶分析也可以只用两种溶液在大约 15 min 内完成，用时远远少于耗时 5～48 h 的引物延伸法、RNase 保护法或涉及放射性标记探针和数种缓冲液的 Northern 分析。

专题 4.4 *LacZ*、SEAP 和 GFP 报告基因分析法

虽然萤光素酶和 CAT 是历史上最常用的报告基因，但也可以用其他几种报告基因。编码 β-半乳糖苷酶（β-gal）的 *E. coli lacZ* 基因就是其中之一（Hall et al. 1983）。β-半乳糖苷酶催化 β-半乳糖苷的水解。用于标准 β-半乳糖苷酶分析的底物包括 5-溴-4-氯-3-吲哚-β-D-半乳糖苷（X-gal）和邻硝基苯-β-D-半乳吡喃糖苷（ONPG），这些底物生成可用分光光度计定量的反应产物。以前，*lacZ* 报告基因经常用作内部对照，用于使用 CAT 或萤光素酶报告基因的转染研究。β-gal 分析的一个局限性是具有内源性哺乳动物酶引起的背景，可以通过选择性地灭活哺乳动物酶的热处理减小这种背景（Young et al. 1993）。β-gal 报告载体可从 Clontech（631707）购买，β-gal 对照载体可从 Promega（E1081）和 Clontech（631709）购买。用于分光光度法检测酶活性的试剂可以从 Promega、stratagene 和 Invitrogen 单独购买，或以试剂盒的形式购买。Roche Applied Science（11758241001）和 Clontech（631712 and 631713）提供用于监测 β-gal 酶活性的敏感化学发光试剂盒，Roche Applied Science 有用于测量 β-gal 蛋白的 ELISA 试剂盒（11539426001）。

另一种有用的报告基因是分泌型人胎盘碱性磷酸酶（SEAP），Clontech 出售的报告质粒（631715 和 631717）上就表达这种酶。使用 Clontech 试剂盒（631704 和 631701），可通过荧光或化学发光检测定量 SEAP 蛋白。该报告基因的一个重要优势是报告蛋白被分泌到培养基中，从而使报告基因检测不需要细胞裂解。因为当测量 SEAP 蛋白表达时，细胞仍维持于培养物中，所以可以监测表达的动力学，并在之后可收集转染的细胞用于其他研究。

目前另一种与萤光素酶几乎一样常用的报告基因是绿色荧光蛋白（GFP），它是从水母 *Aequorea victoria* 中分离得到的分子质量为 28 kDa 的蛋白质（Prasher et al. 1992；Chalfie et al. 1994）。GFP 含有内在的肽生色团，经氧化和紫外光或蓝光激发后可发出绿光（发射峰波长为 509 nm），所发出的光通常在完整细胞中就可以检测，从而产生一种报告基因表达的简单分析方法。使用 GFP 作为报告分子的许多质粒载体（包括稳定和不稳定的形式）可以从 Clontech 获得。从商业化质粒表达的 GFP 含有特定的氨基酸替换，可增强荧光的发射，且有时会改变发射波长和蛋白质的稳定性。稳定的 GFP 有利于弱信号的检测，而不稳定的 GFP 会表现出降低的背景信号，并且对于研究转录的可诱导沉默特别有用。GFP 报告分子通常用于通过荧光显微镜和流式细胞仪监测基因表达。另外，可用荧光计检测 GFP 的荧光。GFP 作为报告基因的主要优势是，尤其是在可使用流式细胞仪的情况下，容易获得有关细胞群体中表达 GFP 报告基因细胞的百分比的详细信息，也容易获得细胞群体中的平均荧光和荧光范围的相关详细信息。该信息在剖析转录控制机制时非常有价值。

间接报告分析的局限性。蛋白质一般比其 mRNA 稳定，因此，需要进行恰当的对照实验以确保在线性范围内测量酶活性。此外，在启动子功能分析过程中，获得的结果可能具有误导性，因为该分析方法不能区分在真正的转录起始位点起始的转录本和那些

在载体或启动子内异常起始的转录本。这些"隐秘"转录本并不少见，在瞬时分析过程中，低水平的隐秘转录总是在质粒内的许多位置上起始。这些转录本的一部分将会翻译成有功能的蛋白质，这些蛋白质将促成酶分析中获得的总信号。通过使用已设计用于阻断从报告基因的异常转录的质粒载体，并通过在初始转染实验中加入合适的对照，可以减弱许多背景（见下文）。

专题 4.5 瞬时转染细胞的分离

在建立瞬时转染分析中，一个关键的挑战是发展有效的细胞转染程序，其效率应足以检测报告基因的活性。如果使用现有的技术未获得有效的转染，增强活性检测的一个策略是将转染细胞从非转染细胞中分离出来，该策略的多种变通方法已用于改进对报告基因的检测。

一种方法是使用由目的控制区调控的报告质粒与第二个质粒共转染细胞，第二个质粒含有在病毒启动子/增强子控制下编码 GFP 或细胞表面蛋白的基因，吸收 DNA 的细胞通常可吸收这两种质粒。可通过荧光激活的细胞分拣系统（FACS）分离表达 GFP 的细胞；可以通过多种方法分离表达细胞表面蛋白的细胞，这些方法包括 FACS 或结合于连接有适当配体或抗体的磁珠。可以从分离的细胞中制备提取物，并检测报告基因的活性。

使用在目的启动子控制下的 GFP 作为报告基因也可以有助于避免由低转染效率引起的问题，因为流式细胞仪可显示出转染细胞群体的平均荧光，同时将未转染细胞的背景荧光作为单独的峰观察。遗憾的是，如果转染率很低（如低于 5%），基于如此少量的细胞中检测到的荧光，难以得出有力的结论。

---

*最小化背景转录并确证起始位点。*最小化背景转录的有用策略是将多聚腺苷酸信号插入紧挨启动子的上游，这样载体内起始的异常转录本将被切割和多聚腺苷酸化。几乎所有从 Promega 得到的报告载体都含有上游多聚腺苷酸化信号。使用该策略，只有产生的稳定的 mRNA 分子是从多聚腺苷酸化下游起始的 mRNA（即启动子区内）。应该注意，该程序将不会消除背景转录本，这些转录本是从插入的启动子内异常的位点上起始的。

如果使用间接分析法得到模棱两可的结果，就可能有必要通过直接分析法确证间接分析的结果，直接分析法包括引物延伸、RNase 保护或 5′ RACE（cDNA 末端的快速扩增）分析等（见第 3 章）；这些直接分析法可提供转录起始位点位置的相关信息。如果通过瞬时转染得到了强报告基因活性，就可能相对易于监测转录起始位点的位置。引物延伸分析和 5′ RACE 通常是应该考虑的首选分析方法。有了这两种方法，可用单个寡核苷酸引物监测来自野生型启动子和一系列启动子突变体的转录。已经为至少两种常见报告基因确定了在引物延伸分析中可产生成功结果的引物序列：萤光素酶 [5′-ACCAACAGTACCGGAATGCC-3′ (Plevy et al. 1997)] 和 CAT [5′-TATATCAACGGTGGTATATCCAGTG-3′ (Liston and Johnston 1999)]。因为与这些引物相杂交的序列位于其报告基因内部，所有这些序列对于多数萤光素酶和 CAT 质粒都是有用的。在

某些情况下，RNase保护分析可能比引物延伸分析更灵敏，但是对于每种启动子/报告基因融合及一些启动子突变体来说可能需要不同的探针。

由于许多细胞系的转染效率低，通常难以直接分析mRNA。为提高灵敏度，可以考虑将细胞群中转染的细胞从未转染的细胞中分离出来的方法（见专题4.5）。可以将增强子置于报告基因的下游，也可以测试具有复制能力的报告质粒（见下文及专题4.6）。使用简单而高灵敏的RT-PCR分析监测mRNA似乎比引物延伸、RNase保护或5′ RACE更有效。遗憾的是，RT-PCR分析（还有Northern印迹）通常不能用于测量瞬时转染分析中合成的特定转录本，这些分析方法测量与探针序列或引物序列互补的所有RNA分子，而不管RNA是否是在正确的起始位点起始。如同上文指出的，瞬时分析过程中大量转录本起始于隐秘的载体和启动子序列。RT-PCR的引物或Northern印迹的探针将与异常转录本及正确起始的转录本一样有效地杂交。

质粒的构建

为了分析启动子的活性，必须将假定的启动子序列插入报告基因的上游。为了选择用于插入报告载体的适当的DNA片段，重要的是要首先确定翻译起始密码子和转录起始位点的位置。对于哺乳动物的基因，转录起始位点通常位于翻译起始密码子上游50～200 bp处。然而，如同第3章中所描述的，简单地假定启动子将在ATG密码子的直接上游是不恰当的，即使通过序列同源性鉴定出潜在的TATA框和其他潜在的调控元件。有可能在目的基因ATG的侧翼序列存在的情况下检测出报告基因的活性，然而，除非已经定位了转录起始位点，否则检测到的活性是由真正的启动子驱动的，还是由遍及整个基因组的隐秘的控制元件驱动的，将仍然不清楚。因为启动子区似乎总是位于转录起始位点周围，所以如果已经清楚地界定了转录起始位点，就会大大增加对结果有效性的信心。

*插入片段的下游边界。*一般来说，为初步研究所选择的启动子片段的下游边界，应在转录起始位点和ATG之间。在这一非翻译前导区中已经报道了功能性启动子元件，因此，尽可能多地包括该非翻译区是有用的（遗憾的是，瞬时转染分析中某些非翻译前导序列可抑制启动子活性。因此，如果通过含有前导序列的质粒未检测到启动子的活性，那么可检验含有较少量前导序列的相关质粒）。如果所使用的报告分析依赖于报告蛋白或酶活性的检测，那么应从插入的启动子片段中省略掉ATG和任何剪接供体序列，因为它们可能干扰报告基因的加工和翻译。或者，来自插入的启动子的ATG可直接融合于报告基因的ATG（即代替报告基因的ATG）。

*插入片段的上游边界。*插入报告载体的启动子片段的适当的上游边界更难预测。某些启动子具有强活性，且只在转录起始位点上游100～300 bp处被适当调控。除非伴随诸如增强子的远程控制区，否则其他启动子就不能被调控或不能产生可检测的活性。因为增强子的常见位置是启动子的上游，所以使用更大的片段可增加检测被恰当调控的活性的可能性。应该注意，在几个功能性启动子区的上游已经检测到组成型活性负调控区（见Malone et al. 1997）。虽然这些区域在内源性基因调控中的生物学意义通常不是未知的，但是它们能干扰启动子活性的检测。

搜索资料文献中相关基因的先例，可洞悉要获得被正确调控的报告基因强活性需要

多长的启动子序列（即序列量），因为属于特定种类的基因之间似乎存在调控机制的相似性。正如第 2 章中所讨论的，为了调控转录，早期应答基因和可诱导的细胞因子基因通常依赖于起始位点的数百个碱基对中的序列（Goodbourn et al. 1985；Treisman 1986；Plevy et al. 1997）。虽然该类基因的转录可能受到更远程控制区的影响，但是通过研究最初的数百个碱基对，可以了解这些基因的大量调控信息。相比之下，某些参与淋巴细胞早期发育的基因的启动子，很多缺少 TATA 盒，并在瞬时转染分析中表现出非常低的或不可检测的活性（Lo et al. 1991）。而诸如 β-珠蛋白基因的其他发育调控基因的启动子，在瞬时转染中相当活跃，但调控性很差（Charnay et al. 1985）。因此，用于该类基因的初始报告质粒可能应包括相对大量的上游 DNA。

*含增强子载体的使用。*当制备初始构体时，建议将启动子片段导入至少两个不同的载体中，一个载体仅含有无启动子的报告基因，而另一个含有位于报告基因下游的组成型激活增强子（如 SV40 增强子）。第一个载体中的插入将揭示所分离的启动子的强度和特性。如果分离到的启动子弱，那么加入到第二个载体中的增强子应该有助于放大信号（如见 Wirth et al. 1987）。包含在标准报告载体的病毒增强子会刺激大多数但非全部启动子的转录（Treisman et al. 1983；Robbins et al. 1986），并且通常不会屏蔽启动子的调控，尽管对每个启动子应以实验论证这种可能性。

*质粒的纯化。*可以通过 CsCl 梯度法（Sambrook and Russell 2001，pp. 1. 65-1. 68）或标准柱层析法（如 QIAGEN QIAfilter Plasmid Maxi Kit）纯化用于转染实验的质粒。对于某些细胞系，DNA 的质量和纯度可大大影响转染效率和报告基因活性。因此，为了优化实验或解决实验中的问题，应测试不同的 DNA 制备方法。如果通过从 DNA 制备物中去除内毒素（如脂多糖）的方法制备质粒，那么就会改善通过某些细胞系获得的结果（如见 Plevy et al. 1997）。QIAGEN 提供的 EndoFree Plasmid Maxi Kit 是含有用于内毒素去除程序的质粒纯化试剂盒。

初步转染实验

*对照转染。*使用含有目的启动子和标准报告基因的报告质粒进行的初步转染实验，应该包括三个主要的对照转染（图 4.7）。一个对照是用无启动子的载体转染，其中载体就是用于插入目的启动子的载体。第二个对照应该使用含有融合于报告基因的强病毒启动子的质粒进行转染。最后，应该用不加质粒或缺少报告基因的质粒进行模拟转染。以强对照启动子检测到的信号将确证细胞被有效转染，而且将提供目的启动子相对强度的相关信息。相对于模拟转染，以无启动子载体检测到的信号指示由载体内隐秘的控制元件引导的报告基因活性。在不存在插入启动子的情况下，每个载体将指导报告基因的转录，但只有一小部分这些转录本胜任翻译。由于这一原因，通常以灵敏的报告基因分析方法检测背景信号。

如果使用含有目的启动子的质粒检测到的信号远比使用无启动子质粒检测到的信号强，那么该差异可能对应真正的启动子活性。如果该信号只是稍强于或者中等程度地强于使用无启动子质粒获得的信号，该信号可能对应真正的（虽然弱）启动子活性，或者来源于在插入片段内起始的异常转录。任何一种情况下，为了确证以酶分析检测到的活性对应于特定启动子活性，进行引物延伸、RNase 保护或 5′ RACE 分析证明转录本起

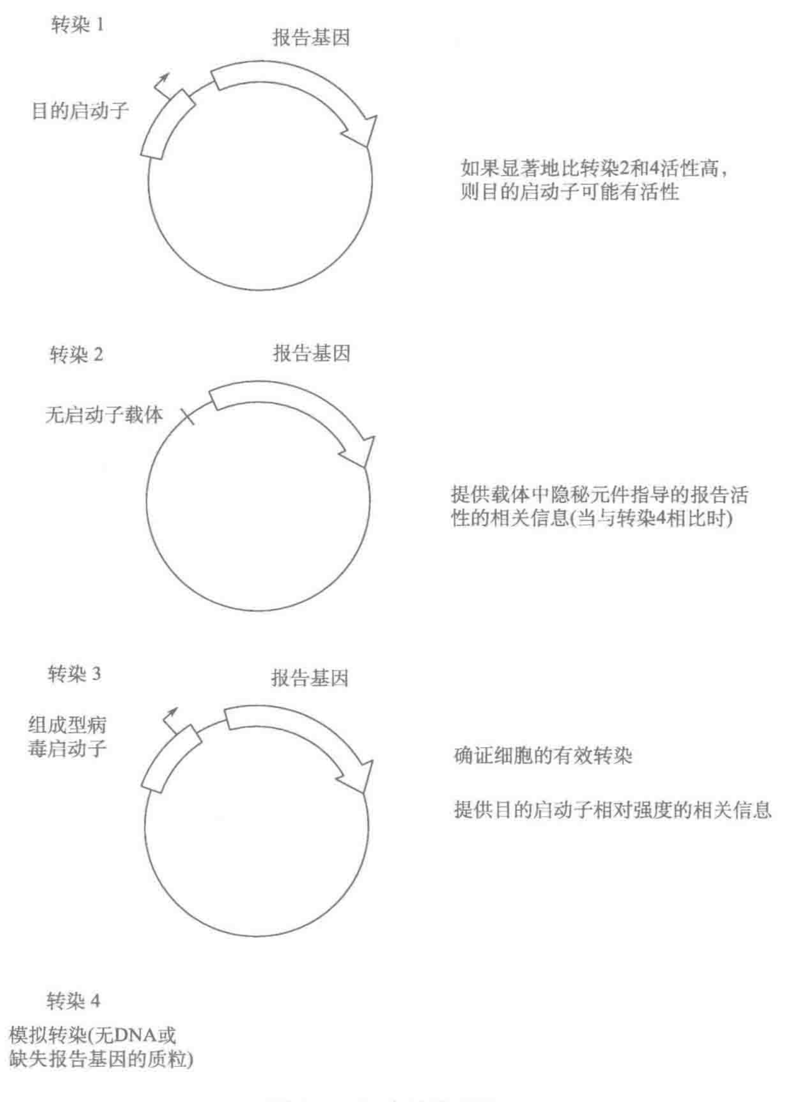


图 4.7 初步转染实验。

始于正确的起始位点是有益的。遗憾的是，如果启动子的活性弱，可能难以实现这种证明。因此，可能难以确证该活性是否对应于真正的启动子活性，除非完成更为广泛的分析。这种广泛的分析应该包括一项实验，即启动子活性受到适当的调控，并依赖于一组不相关联的、功能性的控制元件，正如详细的突变体分析所揭示的那样（第 6 章）。

内对照。当以间接酶分析测量报告基因的表达时，通常包含内对照质粒，以归一化当时所进行的所有转染。该对照是含有标准启动子的质粒，其中，标准启动子融合于不同的报告基因。*Renilla reniformis*（海肾）萤光素酶内对照质粒的 pRL 家族是目前使用最广泛的内对照质粒。当与含有目的启动子的萤火虫（*Photinus pyralis*）萤光素酶报告质粒共同转染时，这些质粒可方便地用作内对照。当使用双萤光素酶分析系统时，

首先在细胞提取物中监测源自萤火虫萤光素酶的发光，然后灭活此信号并通过加入不同的反应底物分析海肾萤光素酶活性（见专题 4.2）。可用的 pRL 对照质粒包含 HSV-TK、SV40 和 CMV 启动子。当比较一系列启动子突变体时，这样的内对照特别有用。然而，确定结果的有效性时主要靠实验的重复性，而内对照并不能减少这种需要。在转染之前和转染程序本身中，分装 DNA 时都会产生测量误差，从而提高不准确归一化的可能性。在内对照质粒中使用的病毒启动子/增强子也有可能竞争极限转录因子（Farr and Roman 1992）。

*进行直接 RNA 分析时所用的对照。*当使用对 mRNA 丰度进行更直接测量的分析方法（如引物延伸、RNase 保护或 5′ RACE）时，可用与上文所描述的对照相似的对照进行初步转染实验。mRNA 测量的优势是：如果得到了明显的特异性信号，就更有把握说明它对应于正确起始的转录。尽管如此，在初步实验期间，重要的是包括模拟转染和无启动子载体的转染，以证明只有当包含启动子时，才能得到预期的条带，因为每种直接分析经常都检测到背景条带。正如上面对酶分析所作的讨论，为了归一化转染效率并有利于 mRNA 制备，应设置对照质粒。然而，因为 mRNA 检测分析涉及很多操作，这些操作会产生样品间的差异，所以所得结果的可靠性必须主要基于实验的几次重复。

评价适当的启动子调控

初步瞬时转染分析的最后步骤是，确定在人工分析中是否适当地调控了研究的启动子。对于细胞类型特异的基因，“适当的调控”是指在与内源基因相同的细胞类型中控制区活跃的程度。对于可诱导的基因，“适当的调控”是指以与内源基因相似的动力学及通过与内源基因相似的试剂诱导控制区的程度。

*可诱导基因。*对于可诱导基因，评价瞬时转染分析对内源基因调控模拟的程度相对简单。使用启动子/报告质粒转染等份的适当细胞，并使细胞从转染程序恢复一段时间(该时间将依特定的情况而改变)，随后，一部分细胞等份用适当的试剂诱导，而其他细胞等份模拟诱导并维持在培养物中，以便同时收集两个细胞群。如果在瞬时分析中启动子得到适当的调控，相对于对照启动子的活性，该启动子的活性应被选择性地诱导。

*细胞类型特异的基因。*作为对启动子细胞类型特异性的粗略测试，可比较表达细胞系和非表达细胞系中的启动子活性。虽然一种或两种细胞系中的微弱或微不足道的活性不能最终证明启动子的活性受到严紧的或正确的调控，但是它提供了调控的初步证据。在大多数情况下，比较多种细胞类型中启动子的活性时会出现两个困难。第一个困难是，在瞬时分析中分析的大多数启动子不像内源基因那样被严紧地调控。通常在某些非表达的细胞类型中检测到瞬时活性，而且瞬时活性不可能在任何细胞类型中完全不存在。这可能归因于一种事实，即在非表达性细胞中可能需要染色质和其他调控区（如沉默子）以保持内源基因无活性。此外，多拷贝质粒的存在可引起泄漏表达，因为部分正、负调控蛋白可能处于极限量。

当比较多种细胞类型中启动子的活性时，遇到的第二个困难是，细胞系间转染效率不同，因而不能直接比较绝对活性。通常使用在强病毒启动子/增强子控制下的报告质粒归一化转染效率，该载体在所有细胞类型中发挥的作用相似，如包含 CMV、SV40 或 HSV-TK 启动子和增强子的对照萤光素酶载体。虽然这种方法被普遍地用于归一化，

但它远非理想，因为难以确定病毒控制区在被检细胞类型中确实有相似的活性。用于比较转染效率的另一种替代方法是进行免疫荧光分析或流式细胞术（如用 GFP 报告质粒转染后），以监测转染细胞的百分比。同样，这也远非理想，因为转染的结果会不同，这不仅与转染的细胞数有关，也与导入每个细胞中的分子数有关。尽管不存在用于比较瞬时分析中细胞系间启动子活性的理想方法，但是推荐用两个或更多个对照质粒归一化启动子的活性。因此，如果一个病毒启动子质粒在一种细胞系中具有异常活性，那么其他对照质粒的活性会显示出不同。遗憾的是，如果观察到了不同，也难以确定哪个病毒启动子最适于归一化使用不同的细胞系获得的信号。

最终，需要更先进的实验确定正在研究的控制区是否真正包含参与细胞类型特异性调控的控制元件。目的是说明特定控制元件中的置换突变影响一部分细胞系（即那些表达目的内源基因的细胞系）中进行瞬时转染的启动子活性，而不影响另一部分细胞系（那些不表达目的内源基因的细胞系）中进行瞬时转染的启动子活性。启动子突变体的差异效应可以提供启动子在不同细胞系中确实受到差异调控的重要证据，以确证差异调控并确定其机制。

通过染色体整合的稳定转染分析

为了发展用于启动子或其他控制区的稳定转染分析，必须考虑的许多问题都与上文对瞬时转染分析所做的讨论相似，即必须选择细胞的来源、转染程序和报告基因，然后构建报告质粒并使用适当的对照进行初步实验，以确定控制区是否正确地起作用。稳定转染分析特有的另一个决策是药物选择方法，这些方法用于分离稳定整合了报告质粒的细胞。

上面的问题和实验方法的一般概要在下文讨论。遗憾的是，关于用稳定转染分析研究基因调控的方法，我们尚未见到可为此提供更详细信息的已发表论文。基因稳定整合后，可用稳定转染过表达某种蛋白质，有关这方面的深入信息，可以查阅 Ausubel 等（1994，单元 9.5）和 Spector 等（1998）。即使有了该信息，也强烈建议首次进行稳定转染实验的研究者依赖于具有稳定转染经验的研究人员的个人建议和辅助。如同下文所讨论的，其中几个步骤，包括抗药性细胞克隆或群体的初步生长和维持，对于初学者来说可能非常乏味和困难。

一般策略

可用两个基本策略进行涉及染色体整合的基本稳定转染分析。最常用的，也是通常首选的方法是用两个基因转染细胞：一个基因含有驱动标准报告基因的目的启动子，而另一个基因含有通常由病毒启动子/增强子控制的抗药性基因。这两个基因可以在同一质粒上，也可以在不同的质粒上。此外，可将边界元件或隔离子元件置于启动子-报告基因盒的两侧，以减少整合位点变化和抗药性基因盒对报告基因转录的影响（见下文；Chung et al. 1997；Zhong and Krangel 1997；Muljo and Schlissel 2002；Hsu et al. 2003；Zhou et al. 2007）。转染后，将细胞短时间孵育（一般 1～2 天），然后使其经受药物选择。获得抗药性细胞克隆或群体后，通过测量报告基因的 mRNA、蛋白质或酶活性，定量启动子活性。

第二个基本策略是将抗药性基因或能够引起细胞转化的癌基因直接置于目的启动子

的控制之下（如见 Fromm and Berg 1982、Zhong and Krangel 1997 和第 5 章），该策略不经常使用。通过该策略得到的克隆数似乎与启动子强度相关。据推测，这是因为抗药性或转化依赖于基因产物的特定阈浓度，即启动子越强，达到阈浓度的可能性将越大。因为克隆数和启动子强度相关，所以可通过对获得的抗药性克隆的简单计数，测量启动子突变体的近似活性。该分析除了相对简单之外，将抗药性基因或癌基因置于目的启动子控制下的优势是克服了第一个策略的局限性。特别是当抗药性基因和报告基因共转染细胞时，它们几乎总是共整合到细胞的同一位置上，即使它们是在不同的质粒上被转染的也是如此，这种共整合几乎总是导致报告基因紧密靠近驱动抗药性基因转录的启动子和（或）增强子。除非报告基因盒两侧有边界元件，否则驱动抗药性基因表达的控制区会影响报告基因的转录。如果将抗药性基因或癌基因直接置于目的启动子的控制下，就不需要另一个控制区。

尽管具有这个潜在优势，但第二个策略具有两个严重的局限性。其一，如果启动子太弱而不能产生足够的基因产物，就得不到克隆。其二，获得的克隆数不可能与启动子强度直接成正比，虽然二者可能相关。因为这些缺点严重地限制了第二个策略的使用，所以下文集中讨论第一个策略。然而，如果对于研究的特定控制区，第二个策略的利大于弊，那么它可提供有关基因调控机制的有价值的信息。

细胞和转染程序

对于稳定分析的细胞和转染程序的考虑与那些对瞬时分析的考虑相似，即细胞必须表达目的基因、必须有可能使细胞在分析过程中一直维持在培养物中，以及必须具有把 DNA 转染进细胞的条件。通常使用瞬时分析建立最佳的转染条件，以优化导致标准报告基因（如萤光素酶或 GFP）有效瞬时表达的条件。重要的是要知道某些转染程序（如 DEAE-葡聚糖转染）对于瞬时转染实验是有用的，但由于长期的毒性问题，对稳定转染却没用（见专题 4.1）。

稳定分析选择细胞的标准与瞬时分析的一个差别是，需要使细胞在培养物中维持较长的时间，细胞也必须经得起药物选择。并非所有的药物选择程序对于所有细胞系都有效，用于致死和选择的最佳药物浓度在细胞系间变化范围很大（见下文）。在确定适当的转染程序时，应注意到，对于以目的控制区进行成功瞬时转染太低的转染效率，通常可在稳定分析中成功。为了使稳定分析取得成功，通常有必要仅使群体中少量细胞吸收 DNA。只要药物选择程序适当，那些经选择仍生长的细胞就是含有转染质粒的细胞克隆。

报告基因和分析方法

为稳定转染选择合适的报告基因时所做的考虑，与为瞬时分析选择报告基因时所做的考虑相似。通过酶分析、荧光或 mRNA 分析可监测的任何标准报告基因，对稳定转染都是合适的报告基因。一个新的考虑是稳定转染分析的使用通常增加有效监测特定转录本的可能性，其中有效监测既可以通过同时使用间接蛋白分析和 mRNA 分析实现，也可以用 mRNA 分析替代间接蛋白分析实现。如同先前所讨论的，某些分析方法提供从特定起始位点起始的转录量的相关信息，这种分析方法的使用可大大增加研究者对结果有效性的信心。

专题 4.6 有复制能力的载体

一种替代的瞬时转染策略是使用含有病毒复制起点的载体，这些复制起点促使质粒复制到高拷贝数（图 4.8）。有些启动子的活性在常规瞬时转染分析中无法检测，而该策略可检测到来自这类启动子的活性。而且，获得的信号可能由于增加的拷贝数而更强，增加的拷贝数使得转录起始位点的监测和确证更容易。该方法的主要缺点是高质粒拷贝数增加了启动子将被异常调控的机会，最大的顾虑是细胞核内重要的转录因子将是极限的。这种情况下，转录可能主要受到能与启动子相互作用的丰度最高的因子的调控，而不是受到对调控内源基因最重要的因子的调控。此外，另一个局限性是病毒复制

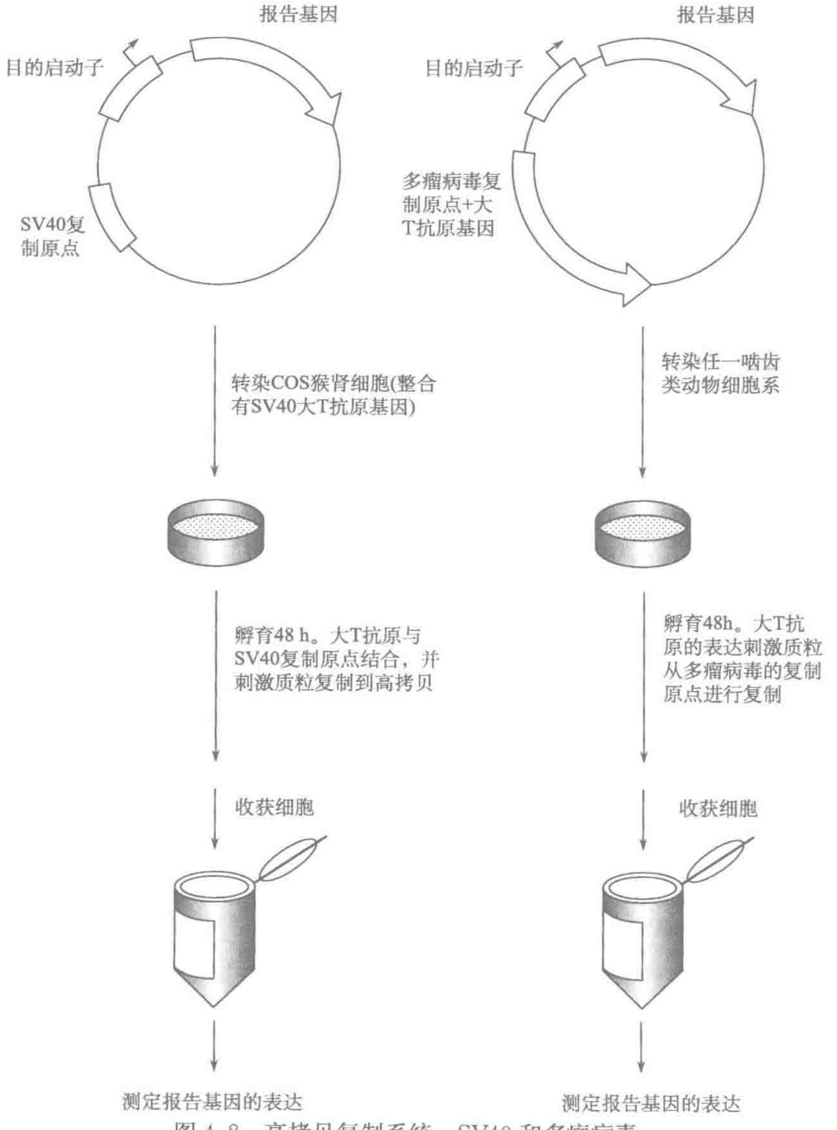


图 4.8 高拷贝复制系统：SV40 和多瘤病毒。

起点的功能通常部分依赖于相邻病毒增强子的存在。因此，几乎所有用于这种分析的质粒都含有病毒增强子，这些增强子可作用于目的启动子，并使不依赖于邻近增强子的启动子活性的研究变得困难。因此，虽然该分析方法对于洞悉参与基因调控的控制区和控制元件可能是有用的，但它远非理想。

最常用于哺乳动物细胞的两个高拷贝复制系统依赖于 SV40 和多瘤病毒复制起点。含有 SV40 复制起点的载体只有在 SV40 大 T 抗原存在的条件下才能在灵长类动物细胞中复制出高拷贝数。众所周知的 COS-7 非洲绿猴细胞系含有稳定整合的编码大 T 抗原基因的 SV40 片段（Gluzman 1981；也见 Sambrook and Russell 2001，p. 11. 114）。人类 HEK293T 细胞也包含一个整合的 SV40 大 T 抗原基因。使用含有 SV40 复制起点的质粒转染 COS-7 或 293T 细胞，经常使质粒复制到 50 000 分子/细胞的拷贝数。含有多瘤病毒复制起点的载体只能在表达多瘤病毒大 T 抗原基因的啮齿动物细胞系中复制。多瘤病毒复制系统已经用于研究多种细胞类型中的转录调控，包括淋巴细胞（Grosschedl and Baltimore 1985；Lo et al. 1991）。为了做到这一点，需要制备既含有多瘤病毒复制起点，又含有多瘤病毒大 T 抗原基因编码序列的质粒（Grosschedl and Baltimore 1985）。可将目的启动子及其相连的报告基因插入同一质粒中。然后，可将质粒转染进任何小鼠细胞系。大 T 抗原将被表达，并与复制起点结合，从而刺激质粒复制到高拷贝数。

抗药性基因和载体

自从 20 世纪 70 年代晚期创立通用方法以来，已有数种选择性基因用于产生稳定的转染体。在最早的研究中，选择性基因能恢复代谢途径中的酶缺陷。例如，胸苷激酶突变型细胞被 HSV-TK 基因转染（Wigler et al. 1977；Mantei et al. 1979），稳定转染 HSV-TK 基因的细胞在 HAT 培养基中被选择，HAT 抑制 TTP 自脱氧尿苷的从头合成，导致对活性胸苷激酶的需要。

从 20 世纪 80 年代早期开始，大量显性选择标记基因的发展摆脱了对突变型细胞系的需要。*E. coli* 黄嘌呤-鸟嘌呤磷酸核糖转移酶（*gpt*，Mulligan and Berg，1981）和氨基糖苷磷酸转移酶（*neo*，Southern and Berg 1982）基因首次用于这一目的。如今，这两个基因仍在普遍使用。*E. coli* GPT 酶催化黄嘌呤经过补救途径向 GMP 的合成。当生长培养基补充氨基蝶呤和霉酚酸阻断哺乳动物酶催化的 GMP 从头合成时，细胞的生存依赖于 *gpt* 基因的整合和表达。*E. coli* 的 *neo* 基因使细胞具有对抗生素 G418 的抗性。当 G418 加入培养基时，细胞的生存依赖于 *neo* 基因的稳定整合及表达。表 4. 1 中列出了用于稳定转染实验的一些其他显性选择标记基因。

确定使用哪种抗药性策略取决于几个因素。第一是获得编码抗药性基因的适当载体，因为有些基因的载体不能从商业来源得到（见下文）。第二是药物的费用，它是变化的，并且如果设计了很多实验，药物的费用就很高，药物的费用受到杀死所用细胞所需药物量的严重影响，而此药物量在细胞系间变动很大。第三是药物使细胞致死和使稳定转染细胞得以选择的速度，现已证明嘌呤霉素在稳定细胞系成功传代的速度和频率两个方面均具有显著效果。最后一个是得到具有特定抗药性方案的经验的研究人员的帮助，有些方案所显示出的特性只有具有该经验的研究人员才能理解。

表 4.1 用于稳定转染哺乳动物细胞的显性选择标记

显性选择标记	基因编码	选择	参考文献	备注
Ble (ZeoR)	印度斯坦链异壁菌（*Streptoalloteichus hindustanus*）ShBle 基因，该酶赋予博来霉素（bleomycin）、腐草霉素（phleomycin）和 Zeocin 抗性	在糖肽类抗生素博来霉素、腐草霉素或腐草霉素衍生物 Zeocin 中选择；所有这三种抗生素都切割 DNA，通过与 *ble* 基因产物的化学当量式结合而得以中和	Mulsant et al. (1988)	已从放线菌（*Actinomyce*）的一个抗性菌株 Tn5 和轮枝链霉菌（*Streptomyces verticillu*）中分离出赋予这些抗生素抗性的其他基因
BSD	土曲霉（*Aspergillus terreus*）*bsd* 基因，该基因赋予杀稻瘟素（blasticidin）抗性	在杀稻瘟素中选择；杀稻瘟素分离自灰色产色链霉菌（*Streptomyces griseochromogenes*），它能够抑制蛋白质的合成	Izumi et al. (1991)	最近作为一种异常快速的选择策略被推销
DHFR	鼠 DHFR，含有赋予 MTX 抗性的单氨基酸突变；DHFR 参与嘌呤的生物合成	在 MTX 中选择；MTX 是野生型 DHFR 的竞争性抑制剂。MTX 不能与突变型 DHFR 结合。因此，表达突变型 DHFR 的细胞对 MTX 有抗性	Simonsen and Levinson (1983)	野生型 DHFR 也可作为选择标记，但细胞必须生长在经过透析的培养基中以除去嘌呤。另外，许多细胞表达内源 DHFR 基因，从而减小了这种选择策略的成功率。野生型 DHFR 通常用于扩增转染的基因
GPT (XGPRT)	*E. coli* 黄嘌呤-鸟嘌呤磷酸核糖转移酶，该酶在无鸟嘌呤培养基中催化从黄嘌呤合成 GMP	在氨基蝶呤和霉酚酸中选择，以阻断 GMP 的从头合成。黄嘌呤和次黄嘌呤也必须提供，作为 XG-PRT 催化合成 GMP 的底物。为了阻止从培养基中的黄嘌呤向 GMP 的合成，必须透析培养基，然后补充胸腺嘧啶和谷氨酰胺。	Mulligan and Berg (1981)	GPT 是发现的第一个显性选择标记，实际上在所有的哺乳动物细胞类型中是有用的
HisD	鼠伤寒沙门氏菌（*Salmonella typhimurium*）HisD，该酶催化从组氨醇合成组氨酸	在组氨醇中选择；组氨醇具有毒性，除非被 HisD 氧化成组氨酸。更严紧的选择是使用在无组氨酸的培养基中的组氨醇；细胞的生存依赖于 HisD 催化的组氨酸合成	Hartman and Mulligan (1988)	使用不含组氨酸的培养基可以提高成功的概率，但是在完全培养基中选择通常成功
HPH	*E. coli* 潮霉素 B（hygromycin B）氨基糖苷磷酸转移酶，该酶赋予潮霉素 B 抗性	在潮霉素 B 中选择；潮霉素 B 阻断蛋白质的合成	Blochlinger and Diggelmann (1984)	当需要两种独立的选择方法时，HPH 选择经常与 Neo 选择联合使用
NeoR	*E. coli* 氨基糖苷磷酸转移酶，该酶赋予新霉素或新霉素类似物 G418 抗性	在 G418 中选择；G418 是一种阻断蛋白质合成的氨基糖苷	Southern and Berg (1982)	G418 的高成本是它的一个缺点

续表

显性选择标记	基因编码	选择	参考文献	备注
OuabainR	大鼠 Na、K-ATP 酶 α1 亚基突变型，赋予乌本苷抗性	在乌本苷中选择；乌本苷是一种强心苷，抑制 Na、K-ATP 酶	Emanuel et al.（1988）	建议主要用于灵长类动物细胞
PAC	白色链球菌（*Streptotnyces alboniger*）*N*-乙酰转移酶，该酶赋予嘌呤霉素抗性	在嘌呤霉素中选择；嘌呤霉素阻断蛋白质合成，除非通过乙酰化脱毒	Vara et al.（1986）	由于嘌呤霉素的一致性好、选择迅速、费用低，通常是首选

DHFR，二氢叶酸还原酶；HisD，组氨醇脱氢酶；MTX，甲氨蝶呤。

可以从多种学术和商业来源获得含有抗药性基因的载体。Promega 提供包含 *hyg*（pGL4. 14 [*luc2*/Hygro] 和 pGL4. 76 [*hRluc*/Hygro]）的 *Photinus* 萤光素酶和 *Renilla* 萤光素酶报告质粒。Clontech 出售包含 *pac*（pPUR）和 *hyg*（pTK-Hyg）的载体。Invitrogen 提供包含 *neo* [pcDNA3. 1（+）]、*hyg* [pcDNA3. 1/Hygro（+）] 以及 *zeo* [pcDNA3. 1/Zeo（+）和 pCMV/Zeo] 的载体。遗憾的是，除了 Promega 的载体，我们不知道其他为启动子或增强子的稳定转染分析特别设计的市售的载体。因此，没有一个质粒含有理想的组件，包括由病毒启动子控制的抗药性基因、位于多克隆位点下游的常见报告基因，以及报告基因序列盒两侧的边界元件。相反，大多数商业载体一般含有位于强病毒启动子和增强子下游的抗药性基因。这些载体可与含有目的启动子的报告质粒共转染细胞。此外，可以将抗药性序列盒切下，并将其插入标准报告质粒。

要确立选择条件，必须选择抗药性策略。之后，检验药物的连续稀释液，以确定致死缺乏整合抗药性基因的细胞所需的药物浓度，这种检验通常在 24 孔或 96 孔板上进行。选择用于实验的浓度通常稍高于致死细胞所需的最小浓度。过高的浓度会减小获得稳定转染的机会。确定了最佳浓度之后，就可以进行初步转染实验。

质粒的构建

两种策略的选择：共转染报告质粒和抗药性质粒，或转染同时含有两个基因的质粒。在同一质粒上包含报告基因和抗药性基因的主要优势是，可以更深入地了解整合的目的控制区和用于驱动抗药性基因转录的病毒启动子/增强子的相对位置。因为病毒启动子/增强子可影响目的启动子的活性，所以使其位于一致的位置上是有益的。然而，大多数细胞都整合多个串联拷贝的质粒，这些串联拷贝的质粒可能以头对头和头对尾的方向连接，这就导致病毒启动子/增强子相对于目的启动子的位置有相当大变化，除非仔细定位插入片段以确定其数目和方向。将报告基因和抗药性基因置于同一质粒上的另一个优势是，基因共整合的频率将会提高；将会分离到相对较少的未整合报告基因的抗药性细胞。相反，如果已经得到存在于不同质粒上的适当的报告基因，就不需要构建新质粒，可以使用不同质粒上的报告基因和抗药性基因立即进行共转染实验。尽管第一种策略具有各种优势，但第二种策略通常也相当有效。

质粒构建应考虑的方面。如果选用共转染策略，那么目的控制区的插入应考虑的方面与用于瞬时转染实验的质粒的构建类似，并可以使用同一载体。然而，强烈建议将边

界元件或绝缘子元件置于报告序列框的两侧，从而使整合位点的变异性和共整合的抗药性序列框的影响最小化（见第 1 章和第 5 章；Chung et al. 1997；Zhong and Krangel 1997；Muljo and Schlissel 2002；Hsu et al. 2003；Zhou et al. 2007）。例如，Muljo 和 Schlissel（2002）描述的载体中鸡 β-珠蛋白绝缘子元件极大地降低了含有小鼠 *Il12b* 增强子和启动子的报告基因序列框在稳定表达中克隆的变异性，也降低了增强子和启动子区染色质结构的变化（Zhou et al. 2007）。该方法的一个有趣的附加说明是分离的 *Il12b* 启动子（即缺少邻近的 *Il12b* 增强子）在两侧含有绝缘子时，其在稳定转染分析中是完全无活性的。当绝缘子不存在时，在生理刺激（如脂多糖）下支持有效转录的稳定克隆可在没有增强子的情况下以低频率获得，这可能是因为与显性选择标记基因有联系的病毒增强子增强了来自 *Il12b* 启动子的转录起始。因此，尽管该启动子在瞬时转染实验中非常有活性，但是在不存在邻近增强子时，该启动子不可用于含有绝缘子的载体的稳定转染分析。当为稳定转染实验构建质粒时，如果对含有促进复制的病毒蛋白的细胞进行稳定转染实验，那么避免用包含来自高拷贝病毒（如 SV40 和多瘤病毒）的复制起点的载体也很重要。活性复制起点将阻止质粒的稳定整合，且源自复制的高质粒拷贝数也将加速细胞死亡。

药物选择

通过最佳的转染条件、致死细胞的最佳药物浓度和适当的质粒，就可以进行初步实验（见 Ausubel et al. 1994，单元 9.5；Spector et al. 1998；Sambrook and Russell 2001，pp. 16.48-16.49）。

*基本的转染和选择程序。*对于每种质粒，要根据已经确立的程序转染一份细胞。转染后，将细胞转移到组织培养皿或多孔板上，这取决于是对抗药性细胞池还是对单个细胞克隆进行分离和分析（见下文）。如果要分离抗药性细胞池，可将细胞转移到标准组织培养皿上。如果分离单个克隆，通常立即等分转染的细胞并将其加到 96 孔板上（有时，用于贴壁细胞的另一种方法是将细胞加入标准组织培养皿中，然后用克隆柱或极限稀释分离单个克隆；见 Spector 等［1998］）。加入各孔的细胞数目将有所不同，这取决于细胞类型、药物选择方案和转染效率，并将需要依经验确定细胞数目。理想情况下，多数孔应最多含有一个抗药性细胞。当细胞自转染程序后恢复 1～2 天时，加入药物开始选择。这时，必须仔细监测细胞，以维持细胞的健康状态，并可长出抗药性克隆。

*药物选择期间细胞的维持。*稳定转染程序中最难的步骤经常是药物选择初期细胞的维持（Ausubel et al. 1994，单元 9.5；Spector et al. 1998）。当被细胞碎片和大量垂死的细胞包围时，抗药性细胞必须保持足以生长的健康状态。同时，它们必须表达足量的由整合的抗药性基因编码的蛋白质，以避免自身死亡。研究人员必须适时传代细胞并更换细胞培养基，这有助于抗性克隆的生长。传代细胞和更换细胞培养基的时机，以及对细胞生长和死亡的一般监测将因细胞系的不同而不同。许多情况下，决定稳定转染实验成功的最重要的因素是作出主观判断的研究人员的经验。鉴于此，强烈建议首次发展稳定转染分析的研究人员获取其他人的辅助和建议，这些人具有进行稳定转染程序的经验，并能在药物选择期间凭视觉检查转染细胞。

*分析细胞克隆和细胞池的选择。*所推荐的用于分析控制区活性的策略是，分离并分

析由每个报告质粒获得的数个单个克隆，以确定活性范围。正如早先所描述的，没有LCR或绝缘子的控制区的活性在不同稳定细胞系间通常变化相当大，因为整合位点的性质可引起控制区活性的增强或抑制。此外，整合质粒的拷贝数也有所不同，并影响控制区的活性。

通过分析用特定报告质粒得到的数个独立克隆，可以获得所得到的活性范围和平均活性的相关知识。当尝试比较控制区在不同细胞系中的活性，或尝试比较野生型控制区和突变型控制区时，这一信息非常有价值。克隆与克隆之间的变异性使这些对比难以进行，但是，如果知道了变异性的程度就能很有帮助。如果目的是监测稳定转染细胞中特定控制区的诱导活性，那么分离和分析单个细胞克隆就同样重要，这是因为整合位点会影响诱导的程度。

*分析细胞克隆。*尽管对几个不同克隆的分析是有价值的，但是在多数情况下，没有必要严格地确定每个假定克隆都确实是克隆的。如果96孔板上的一小部分孔中含有抗药性群落，那么每个群落就可能是克隆的。对于某些实验，明确地确定每个群落都是克隆的（通过极限稀释对细胞进一步亚克隆，以及通过Southern印迹分析确定整合位点）可能很重要。然而，因为几个细胞克隆将会被比较，并且预计由于整合位点的差异有相当大的变化，因此某些样品的寡克隆一般不会产生大的影响。拷贝数的严格确定（通过定量的Southern印迹或实时PCR）也可能没有必要，除非需要精确确定以单个克隆观察到的变化中有多少变化归因于拷贝数的变化，而又有多少变化归因于整合位点的变化。在某些情况下，该信息是有用的，如确定一个控制元件是否符合LCR的定义。然而，对于大多数基因调控研究来说，该信息价值有限。

*分析细胞池。*分离几个单个克隆的替代方法是比较几个独立的细胞池，这一策略的潜在优势是由整合位点和拷贝数的差异造成的各池中的变异可被平均，从而减小池间的差异；其缺点是如果一个或少数克隆表现出由报告基因或抗药性基因的整合位点或表达水平造成的生长优势，池中每个克隆的相对丰度会随时间剧烈变化。其实，稳定转染细胞的细胞池可以非常迅速地成为寡克隆或单克隆。因此，不能假定对细胞池的分析可以产生对许多克隆系进行单独分析时获得的平均结果。而且，在某一时刻以细胞池获得的结果可能不同于将细胞传代数周后获得的结果，这是因为细胞的选择性过度生长十分常见。因此，需要注意的是，相对于当分离了单个克隆时检验的数目，细胞池的使用不能减少以每个报告质粒所需检验的样品数目。虽然细胞池的使用理论上可产生来自大量单个克隆的平均报告基因信号，但是一个或少数几个克隆选择性过度生长的可能性使得对几个细胞池进行检验成为必要。

对照及结果的解释

稳定转染分析中，需要用对照确证启动子受到正确调控。如果启动子预期是细胞类型特异性的，那么用不同细胞系进行稳定转染实验就可以确定这一点。正如瞬时转染实验，用病毒启动子/增强子驱动的报告基因进行平行实验，将有利于对从不同细胞系获得的结果进行归一化处理。考虑到，用对照启动子/增强子获得的活性范围提供细胞系间关于平均拷贝数与整合位点可及性的内在差异的相关信息，如果启动子是可诱导的，通过用不同启动子控制的报告基因进行相似的实验，从而证明诱导是目的启动子特异

的；如果转录只是从目的启动子被诱导，则该结果确证的是启动子内的特定控制元件负责这种诱导。

技术

哺乳动物细胞的常用转染方法

诱导培养的哺乳动物细胞吸收外源 DNA 的最早的方法发展于 20 世纪 60 年代晚期和 70 年代早期（McCutchan and Pagano 1968；Graham and van der Eb 1973）。这些最初的细胞转染方法使用 DEAE-葡聚糖或磷酸钙将 DNA 肿瘤病毒的野生型和突变型基因组导入培养细胞，继而使细胞表达病毒蛋白质并产生感染性病毒颗粒。这些方法很快被用于转染包含特定基因或目的控制区的质粒。

普遍使用的两种常规转染实验是：瞬时转染和稳定转染。在瞬时转染实验中，质粒只存在于一代或两代细胞中，这是因为质粒或者不能复制，或者复制出非常高的拷贝数而具有毒性。基因表达发生在附加型质粒上，且必须在 72 h 这一相对较短的时段内监测。通常需要用灵敏的分析方法监测基因的表达，因为对于大多数细胞系，都只有一小部分细胞能吸收外源 DNA。

在稳定转染实验中，外源 DNA 被维持许多细胞世代。通常通过选择显性可选择标记基因的表达，从而实现外源 DNA 在细胞中的长期维持。当外源 DNA 整合到染色体上时，标记基因稳定表达；或者是某些病毒复制起点使附加体长期维持。在这两种情况下，选择过程将产生一个细胞群，该细胞群中的所有细胞都含有外源 DNA。

转染实验最常用于两个目的。第一，可将目的转录控制区插入质粒，并用于调控报告基因的表达，以监测控制区的活性。第二，可将编码目的蛋白的 cDNA 置于启动子和增强子的控制下，引起蛋白质的表达或过表达。转染实验也用于研究许多其他基本过程，包括 DNA 复制、重组和 RNA 加工。

已有很多方法用于转染哺乳动物细胞（见专题 4.1）。4 种最常见的方法是磷酸钙转染法、DEAE-葡聚糖转染法、电穿孔法和脂质体转染法。下文给出了前三种方法的基本实例（不包括脂质体转染法，因为可从公司买到大量不同脂质体转染的试剂，每种试剂都有公司建议的特定转染条件）。下文描述的每种方案都是为指定的细胞系或细胞类型专门设计的，并只对方案作一般介绍。转染参数在细胞系间有所不同，因此，在用目的细胞系开始实验之前，应查阅已发表的文献，以确定已取得成功的转染条件。

这里提供的方案针对很少或没有接触过这一技术的研究人员，只是对转染实验中涉及的步骤做基本的介绍。有关磷酸钙、DEAE-葡聚糖和电穿孔法转染的优化和操作的更为详细的信息，见 Ausubel 等（1994，第 9 章）、Spector 等（1998），以及 Sambrook 和 Russell（2001，第 16 章）。对于不熟悉转染方法的研究人员，强烈推荐这些参考资料。不管选择哪种转染方法，应该改变多种参数，优化转染效率。即使文章中已经报道了相关参数，在新的实验室优化参数也是有价值的。

尽管这些方法最为常用，但对于难以转染的细胞系可以尝试几种其他方法（见专题

4.1)。对于用这三种基本程序难以转染的大多数细胞系，推荐测试少数几种市售的脂质体转染试剂。这些专利试剂大多数都比较相似，但是每种试剂都具有可用特定细胞类型成功转染的独特结构特征。遗憾的是，难以预测哪种试剂最可能对给定的细胞系有效。

方案 4.1　3T3 成纤维细胞的磷酸钙转染

时线与安排

这里描述的用于鼠源 3T3 成纤维细胞转染的磷酸钙转染方案来源于经 Parker 和 Stark（1979）修改后的 Graham 和 van der Eb（1973）的方法。通常细胞在 60 mm 或 100 mm 组织培养皿中生长，一般经过 1 天后转染；转染时细胞大约呈 50%融合，这使大多数细胞在不直接与其他细胞接触的情况下被转染，这也使细胞在转染后约 48 h 达到 100%融合度，即收集大多数瞬时转染的细胞时使细胞达到 100%铺板。

该程序首先形成含有 DNA 和磷酸钙的沉淀。通过将含有磷酸盐的缓冲液（HEBS）与含有 DNA 和氯化钙的溶液逐渐混合，形成沉淀，然后将该混合物铺到含有贴壁培养的细胞和生长培养基的平板上。DNA-磷酸钙沉淀附着在细胞上，并通过细胞吞噬作用被吸收。细胞暴露于沉淀一定时间后，去除含有沉淀的培养基，并加入新鲜的培养基，以预防毒性。对于该程序，需要凭经验优化 DNA 的用量和在磷酸钙沉淀中暴露的时间，此外，如果需要，也要依经验优化 DMSO 或甘油的处理条件。

第 1 天：形成 DNA-磷酸钙沉淀并加入到细胞中。

第 1 或第 2 天：去除沉淀并加入新鲜生长培养基。

第 2 天或第 3 天：对于瞬时转染，收集细胞；对于稳定转染，加入药物开始选择。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

$CaCl_2$（1 mol/L）〈!〉

　用 0.45 μm 的硝化纤维素膜过滤灭菌。－20℃保存。

细胞，3T3 鼠成纤维细胞

　转染时细胞有 50%的融合度

目的质粒 DNA

　为灭菌 DNA，DNA 可用乙醇沉淀并重悬于灭菌水中。然而，对于瞬时转染实验，灭菌通常并非必要。

生长培养基（如含有 10%胎牛血清的 DMEM）

2×HEPES 缓冲盐溶液

　　12 mmol/L 葡萄糖

　　50 mmol/L HEPES

10 mmol/L KCl〈!〉
280 mmol/L NaCl
1.5 mmol/L $Na_2HPO_4 \cdot 2H_2O$〈!〉

用 10 mol/L NaOH 调节 pH 至 7.05；准确的 pH 对有效转染很关键。用 0.45 μm硝酸纤维素膜过滤灭菌。−20℃保存。

首选的选择药物（可选）

设备

平板（培养，60 mm 或 100 mm）
培养箱，设定在 37℃
试管（6 ml，透明，灭菌）

程序

步骤 1：配制缓冲液

配制 2×HEPES 缓冲盐溶液（HEBS）和 1 mol/L $CaCl_2$。

步骤 2：DNA 沉淀和转染

下面指定的量是用于转染 60 mm 平板上的细胞。对于 100 mm 平板，溶液量应加倍或增为 3 倍。

1. 按下述反应体系将试剂依次加入 6 ml 试管。

H_2O	(187.5−*X*) μl
总 DNA（20 μg）	*X* μl
1 mol/L $CaCl_2$	62.5 μl
总量（每个转染）	250 μl

注意：对于 60 mm 平板，总 DNA 的指定用量是 20 μg，尽管某些研究者更喜欢使用多一些或少一些。对于许多实验，最好使用 20 μg 目的质粒。然而，较少量的质粒可以和足量的"填充"DNA（通常为空质粒载体）结合使用，以达到 20 μg 的总量。在其他实验中，20 μg 将包括特定量的报告质粒及其合并使用的转录因子表达质粒。

为了灭菌 DNA，DNA 可用乙醇沉淀并重悬于灭菌水中。但是，在瞬时转染实验中，灭菌通常并非必要。

2. 往管中滴加 250 μl 2×HEBS。每加入一滴，用力轻弹管壁或涡旋。室温静置 20～30 min，使沉淀形成。溶液应成为均一雾状，且在显微镜下看到细小沉淀。

3. 将沉淀加入盛有 5 ml 生长培养基（通常是 DMEM/10% FBS）、生长 50%融合度 3T3 细胞的 60 mm 平板中。

注意：沉淀体积与培养基体积的比值应大约为 1∶10。

加入沉淀的另一种方法是，首先从平板中吸出培养基，然后加入沉淀，再倾斜平板以保证完全覆盖，并在室温下孵育 30 min。加入 5 ml 生长培养基。某些研究者发现该方法能够显著提高转染效率。

4. 将细胞于 37℃孵育 6～8 h 或过夜（最长 12 h）。

注意：每种细胞系对沉淀有不同的耐受力。对于大多数细胞系，过夜孵育是可以接受的，但在存在沉淀的情况下应不超过 12 h。

5. 轻轻吸出培养基，加入 5 ml 新鲜培养基。

6. 适当处理细胞：

- 对于瞬间转染实验，加入沉淀约 48 h 后收集细胞。
- 对于稳定转染实验，在加入沉淀后的 1～2 天，加入用于细胞选择的药物。

疑难解答

可能原因：低转染效率

解决方案：加入沉淀的另一种方法是，先从平板中吸出培养基。加入沉淀，倾斜平板以保证完全覆盖，并在室温下孵育 30 min，加入 5 ml 生长培养基，某些研究者发现该方法能够显著提高转染效率。另一种可以提高磷酸钙转染效率的处理是用 DMSO 或甘油短暂“休克”，认为这些化学药品具有增强细胞膜通透性的作用（见 Ausubel et al. 1994，单元 9.3；Sambrook and Russell 2001，第 16 章）。

方案 4.2　淋巴细胞系的 DEAE-葡聚糖转染

时线与安排

这里所描述的对于 B 淋巴细胞系和 T 淋巴细胞系的 DEAE-葡聚糖转染法源自 Luthman 和 Magnusson（1983）、Lenardo 等（1987）的方法。对于每个转染实验，大约使用 10^7 指数期生长的细胞。所描述的程序适用于非贴壁的悬浮生长的淋巴细胞。有关适用于贴壁细胞的几个 DEAE-葡聚糖转染方案见 Ausubel 等（1994）、Sambrook 和 Russell（2001，第 16 章）。

DEAE-葡聚糖常用于瞬时转染，但由于其具有毒性，很少用于稳定转染。这里所描述的 DEAE-葡聚糖法，首先用 DNA/DEAE-葡聚糖混合物孵育，然后在加有氯喹的生长培养基中第二次孵育。带有负电荷的 DNA 与 DEAE-葡聚糖聚合物上带正电荷的 DEAE 结合，然后细胞通过胞吞作用吸收该复合物。其中，氯喹通过中和可降解 DNA 的溶酶体水解酶提高转染效率。DEAE-葡聚糖法的一个显著优势是，每次实验需要较小量的 DNA。

第 1 天：先后用 DNA/DEAE-葡聚糖混合物处理细胞，再用氯喹处理细胞。

第 3 天：收集细胞。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

$CaCl_2/MgCl_2$ 原液：
  10 mg/ml $CaCl_2$〈!〉
  10 mg/ml $MgCl_2$〈!〉
  通过 0.45 μm 硝酸纤维素膜过滤灭菌，4℃保存。
细胞，淋巴细胞，悬浮培养（约 10^6 细胞/ml）
氯喹（10 mmol/L）（100×原液，用水配制）〈!〉
  通过 0.45 μm 硝酸纤维素膜过滤灭菌。分装成 5 ml 等份，−20℃保存。
DEAE-葡聚糖原液（10 mg/ml，用 TD 缓冲液配制）〈!〉
  通过 0.45 μm 硝酸纤维素膜过滤灭菌，4℃保存。
目的质粒 DNA
生长培养基（如含 10%胎牛血清的 RPMI）
TD 缓冲液

试剂	g/L
NaCl	8.0
KCl〈!〉	0.38
$Na_2HPO_4 \cdot 7H_2O$〈!〉	0.1
Tris	3.0

调节 pH 至 7.4。高压灭菌。

设备

离心机（临床）
T 细胞培养瓶（75 ml）
培养箱（37℃）
试管（锥形，15 ml，灭菌）

程序

第 1 步：配制缓冲液

配制 TD、$CaCl_2/MgCl_2$ 原液、DEAE-葡聚糖原液和氯喹原液（100×）。

步骤 2：转染

1. 配制新鲜 TS 缓冲液：混合 100 ml TD 缓冲液和 1 ml $CaCl_2/MgCl_2$ 原液。

2. 配制 DNA/DEAE-葡聚糖混合液：

TS 缓冲液	950 μl
质粒 DNA	2 μg
DEAE-葡聚糖原液	50 μl

3. 在台式离心机中，2000 r/min 离心 5 min，沉淀 10^7 个细胞（通常用约 10 ml 密度为 10^6/ml 的细胞培养物）。用 15 ml 灭菌一次性锥形管，倒掉培养基。

4. 将细胞悬浮于 5 ml TS 缓冲液中，以洗涤细胞。如上离心，沉淀细胞，并弃去上清液。

5. 将细胞悬浮于 TS/DNA/DEAE-葡聚糖混合液中（1 ml）。

6. 在锥形管中于室温下孵育 15 min。

7. 加入 10 ml 补充有 100 μmol/L 氯喹（对 10 mmol/L 原液按 1∶100 比例稀释）的含血清培养基（对淋巴细胞通常是 RPMI 和 10% FBS)。

8. 在锥形管中于 37℃孵育 30 min，将管侧置于孵育器中。

9. 2000 r/min 离心 5 min，沉淀细胞，弃上清液。

10. 用 5 ml 加血清生长培养基洗涤细胞。

11. 将细胞悬浮于 5 ml 加血清生长培养基中，并将其转移到含有 25 ml 加血清生长培养基的 T 细胞培养瓶（T75）中。

12. 37℃孵育，大约 48 h 后收集细胞。

疑难解答

可能原因：转染效率低。

解决方案：用 DMSO 或甘油处理增加细胞膜的通透性，可提高 DEAE-葡聚糖转染的效率。DNA 的用量、DEAE-葡聚糖的浓度及用 DEAE-葡聚糖处理的时间均应优化。而且，也应在存在和不存在氯喹、用和不用甘油或 DMSO 休克的情况下，测试 DEAE-葡聚糖转染方法的有效性。也可以改变氯喹处理和甘油/DMSO 休克时间。对于某些细胞系，最适条件可能包括相对短时间的 DEAE-葡聚糖处理和长时间的氯喹处理。因为两种化学药品均对细胞有毒，所以如果一种化学药品的处理时间减为最短，细胞就能够耐受另一种化学药品的长时间处理。

方案 4.3　RAW264.7 巨噬细胞的电穿孔转染

时线与安排

将 DNA 导入培养的真核细胞的电穿孔法最初是由 Neumann 等（1982）发展的，这里所描述的用于 RAW264.7 巨噬细胞的方法源自 Plevy 等（1997）。

电穿孔装置可以从数个公司获得，并被分为两种基本类型（Potter 1988）。大多数装置都提供通过 DNA 样品的指数衰减的脉冲电流。电流是由放电之前充电到确定电压的电容器发出的。可改变电容器的大小及其充电电压，以优化各个细胞系的转染效率（增大电容器的大小可增加特定电压下储存的电流量，从而增加脉冲时间）。第三个重要

的变量是含有细胞和 DNA 的缓冲液的电阻。第二类电穿孔装置提供方波电流和重复脉冲的电势，二者均有利于某些细胞系的转染。Bio-Rad Gene Pulser Xcell（165-2661）可使用指数波或方波形。

该实验首先将 DNA 和待转染的细胞置于可在电击期间产生电阻的适当溶液中，然后将溶液置于与所用电穿孔装置兼容的比色杯中，再对细胞短暂电击。在电击前后的短时间内，将比色杯在室温下孵育（对于某些细胞系，在冰上孵育）。最后，将细胞连同生长培养基和血清转移至组织培养平板或 T 细胞培养瓶中。电穿孔法的一个有吸引力的特点是更为迅速，且比其他方法涉及较少的操作。

第 1 天：在 DNA 存在的情况下电击细胞。

第 2 或第 3 天：对于瞬时转染，收集细胞；对于稳定转染，加入药物开始选择。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

细胞（巨噬细胞，RAW264.7）

当细胞生长到 50%～80%融合度时，收集细胞用于转染。

目的质粒 DNA [约 400 μg/ml，在磷酸缓冲盐（PBS）中]

生长培养基（如含有 10%低内毒素胎牛血清的 DMEM）

磷酸缓冲盐（PBS）

试剂	10×PBS 原液/(g/L)	1×PBS 工作液/(mmol/L)
KCl〈!〉	2.0	2.7
KH_2PO_4〈!〉	2.0	1.4
$Na_2HPO_4 \cdot 7H_2O$〈!〉	11.5	4.3
NaCl	80.0	137.0

调节工作液的 pH 至 7.3。

首选的选择药物（可选，见步骤 12）

设备

离心机（临床）

比色杯（电穿孔，0.4 cm 电极间距；来自如 Gene Pulser/MicroPulser 比色杯；Bio-Rad，165-2088）

电穿孔装置（如 Gene Pulser Xcell 真核系统；Bio-Rad，165-2661）

Bio-Rad Gene Pulser Xcell 能够使用指数波形和方波形。

血细胞计数器

培养箱（37℃，加湿）

培养皿（60 mm 和 100 mm 或 150mm）

试管（锥形，15 ml 和 50ml）

程序

步骤 1：配制缓冲液

步骤 2：转染

1. 在 100 mm 或 150 mm 培养皿上，于生长培养基中培养 RAW264.7 巨噬细胞。当细胞融合度处于 50%～80%时，从平板上刮取细胞用于转染，将其连同生长培养基转移至 50 ml 一次性锥形管中。

2. 用血细胞计数器测定细胞密度。

3. 在 50 ml 试管中，4℃下 1500 r/min 离心 10 min，沉淀细胞，弃上清液。

4. 以 3.75×10^7 细胞/ml 的浓度将细胞重悬于生长培养基中。

5. 向电穿孔比色杯中加入 200 μl 细胞和 20 μg 质粒 DNA（50 μl 体积，在 1×PBS 中）。

6. 将比色杯室温孵育 10 min。

7. 使用 Bio-Rad Gene Pulser，用 960 μF 电容器，施加 250 V 电压电击每个比色杯。

注意：应间歇地确定最佳电压和电容，因为它们会随着细胞的健康状态和年龄有所变化。此外，使用 Bio-Rad Gene Pulser II 优化 RAW264.7 细胞转染的电穿孔条件。Bio-Rad Gene Pulser II 依靠指数衰减，是多年来市场上主要的电穿孔装置。近来，Bio-Rad 中断生产该装置，代之以可提供指数波形或方波形的 Gene Pulser Xcell。

8. 将比色杯于室温下孵育 10 min。

9. 将细胞转移至含有 5 ml 1×PBS 的 15 ml 一次性锥形管中，洗涤细胞。

10. 4℃下 1500 r/min 离心 5 min，沉淀细胞。吸出上清液。

11. 将细胞重悬于 5 ml 生长培养基中，将其转移至 60 mm 培养皿，在 37℃、适宜的湿度条件下孵育。

12. 适当处理细胞：

- 对于瞬时转染，转染后约 48 h 收集细胞。
- 对于稳定转染，转染后 1～2 天加入用于细胞选择的药物。

疑难解答

可能原因：转染效率低。

解决方案：电穿孔法难以优化，因为文献中已报道了许多种缓冲液。缓冲液对于电穿孔法是一个关键参数，因为它界定电击期间的电阻。Ausubel 等（1994）建议，在程序的优化期间可测试 4 种标准缓冲溶液。对于每种缓冲液，必须优化电压和电容。

方案 4.1～4.3 的附加说明

1. 在上文描述的 DEAE-葡聚糖方案中包括氯喹处理。氯喹通过中和可降解 DNA 的溶酶体水解酶提高某些（但非所有）细胞系的 DEAE-葡聚糖转染效率。可提高 DEAE-葡聚糖或磷酸钙转染效率的另一种处理是用 DMSO 或甘油短暂“休克”。据认为，这些化学药品通过增加细胞膜的通透性而发挥作用（见 Ausubel et al. 1994，单元 9.3 或 Sambrook and Russell 2001，第 16 章）。

2. 除了上文描述的可选处理外，应该改变多个参数优化转染效率。即使参数已在文献中有所报道，在新实验室中优化也是值得的。对于磷酸钙程序，要优化的关键参数是 DNA 用量、暴露于磷酸钙沉淀的时间和用 DMSO 或甘油的选择性处理。对于 DEAE-葡聚糖法，应优化 DNA 用量、DEAE-葡聚糖浓度和 DEAE-葡聚糖处理的时间。此外，还应在存在和不存在氯喹、用和不用甘油或 DMSO 休克的情况下测试 DEAE-葡聚糖法。也可改变氯喹处理和甘油/DMSO 休克的时间。对于某些细胞系，最佳条件可能包括相对短时间的 DEAE-葡聚糖处理和长时间的氯喹处理。因为这两种化学药品均对细胞具有毒性，所以如果一种化学药品的处理时间减为最短，细胞就可以耐受另一种化学药品的长时间处理。电穿孔法难以优化，因为文献中不断报道多种缓冲液。缓冲液是电穿孔法的关键参数，因为缓冲液界定电击期间的电阻。Ausubel 等（1994）建议，在程序的优化期间可检测 4 种标准缓冲溶液。对于每种缓冲液，都必须优化电压和电容。

3. 虽然上文描述的三种转染方法都是应用较为广泛的，但是对于用这些方法难以转染的细胞系，可以尝试用其他几种方法，尤其是脂质体转染法，专题 4.1 简述了多种方法。对于用这三个基本程序难以转染的大多数细胞系，推荐测试少数几个商业来源的脂质体转染试剂，这些专利试剂大多数都彼此相似，但是每一种都具有可能使特定细胞类型成功转染的独特结构特征。遗憾的是，难以预测哪种试剂最有可能对哪些特定细胞系起作用。

技术

常用报告酶分析法

当对启动子发展瞬时或稳定转染分析时，它的主要目的是定量启动子的强度。启动子强度一般与 RNA 聚合酶合成的特定 mRNA 转录本数目成正比。为了区分源自转染启动子的转录本和源自内源基因的转录本，通常将启动子融合到异源基因的编码序列上。用于定量异源基因转录的相对直接的方法包括用核连缀分析测定从头 RNA 合成（第 2 章），或通过引物延伸、RNA 酶保护或 Northern 印迹法测定稳态 mRNA 水平（见第 3 章）。遗憾的是，如果细胞的转染效率低，在瞬时转染实验期间，可能难以通过这些方法检测 mRNA。如果启动子弱，在稳定转染分析中也可能难以检测 RNA。由于这些问题，启动子通常融合于编码可用高灵敏性分析定量的酶的异源报

告基因。转染细胞群体中的报告蛋白的活性或荧光与稳态 mRNA 水平大约成正比。报告蛋白分析的其他优点包括简单和快速。

如专题 4.2～专题 4.4 中所描述的，已有数种不同的报告基因和报告分析用于定量转录。三种最常见的报告分析是蛍光素酶（专题 4.2)、CAT（专题 4.3）和 β-半乳糖苷酶（专题 4.4)。第 4 种常见的报告基因编码 GFP，GFP 已经逐渐广泛地用于报告分析，因为流式细胞仪已经变得较为常见。下文给出了前三种分析的简要方案，这些分析及其他报告分析的其他信息，见 Ausubel 等（1994，单元 9)、Sambrook 和 Russell (2001，pp. 17.30-17.51) 及 Promega (1996)。

方案 4.4　萤光素酶分析

时线与安排

这里所描述的基本的萤火虫萤光素酶分析是 de Wet 等（1987）和 Promega（1996）所描述的方法的修改版。该程序使用 Promega 所售的溶液，Promega 是萤光素酶产品的主要供应商。虽然使用在实验室中配制的溶液可以有效地进行该分析（见 Ausubel et al. 1994，单元 9)，但是 Promega 的溶液方便又相对便宜，并且包括产生较强信号的专有修改。Promega 也声称其溶液通过增强酶的周转可延长萤光信号。

在该分析中，用含有去污剂的缓冲液裂解转染萤光素酶报告质粒的细胞，或者，通过冻融裂解细胞（见方案 4.5)，然后通过微量离心去除细胞碎片。在大多数情况下，用光度计测量萤光素酶的活性。某些光度计直接向细胞裂解物中注入含有 D-萤光素底物、ATP 和辅酶 A 的溶液，然后测量产生的荧光。光度计的自动化使信号在注入瞬间就可被测量，这可以增加结果的一致性，因为混合后荧光会相当快地减弱。对于手动光度计，将底物溶液和细胞裂解物手动混合，并在混合后确定的时间（通常为 10 s）读取荧光值，也可以在闪烁计数计中测量萤光素酶活性（见 Nguyen et al. 1988 和 Promega 1996)。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

细胞，用目的质粒和萤光素酶报告基因转染

双萤光素酶报告基因分析系统（Promega，E1910)（可选)

萤光素酶分析试剂（Promega，E1483)

或者，在实验室中准备含有 D-萤光素底物、ATP 和辅酶 A 的分析试剂（Ausubel et al. 1994，单元 9)。

磷酸缓冲盐（PBS)

试剂	10×PBS 原液/(g/L)	1×PBS 工作液/(mmol/L)
KCl 〈!〉	2.0	2.7
KH_2PO_4 〈!〉	2.0	1.4
$Na_2HPO_4 \cdot 7H_2O$ 〈!〉	11.5	4.3
NaCl	80.0	137.0

调节工作液的 pH 至 7.3。

蛋白质分析试剂（如 Bradford）

5×报告基因裂解缓冲液（Promega，E3971）

设备

离心机（临床）（可选，见“对于非贴壁细胞”，步骤 1：2）

光度计

微量离心机

自动移液器

试管（锥形，15 ml）（可选）

试管（离心，1.5 ml）

程序

步骤 1：收集并裂解细胞

1. 去除转染细胞的生长培养基。
2. 收集并裂解细胞。

注意：在裂解缓冲液中用去污剂裂解细胞。也可以通过冻融或者用实验室制备的含有 TritonX-100 的缓冲液裂解细胞。

对于贴壁细胞

a. 在培养皿上直接用 PBS 洗涤贴壁细胞。

b. 对于每个 60 mm 细胞培养皿，直接加入 400 μl 的报告基因裂解缓冲液。

c. 将细胞从平板上刮下，并将其转移至离心管中。

对于非贴壁细胞

a. 将非贴壁细胞转移到 15 ml 管中，1500 r/min 离心 5 min，用 PBS 洗涤细胞。

b. 每 10^6 个非贴壁细胞，需向 15 ml 管中加入 100 μl 报告基因裂解缓冲液。

c. 用自动移液器重悬细胞，促进细胞团块解体。

d. 将裂解的细胞转移到离心管中。

3. 在冰上将细胞在裂解缓冲液中孵育 10～20 min，使细胞完全裂解。
4. 用微量离心机以全速离心 5～10 min，沉淀细胞碎片。

注意：不充分的离心将导致不一致的光度计读数。

5. 进行 Bradford 分析，以确定各裂解物中总蛋白浓度。

步骤 2：荧光素酶分析

进行该分析时，样品和分析试剂均应置于室温。建议做双重分析，以确保结果的有效性。

1. 裂解后 1 h 内，将 20 μl 上清液和 100 μl 荧光素酶分析试剂混合。

注意：如果 *Renilla* 荧光素酶质粒作为内对照包括在转染实验中，也可以根据生产厂家的说明使用 Promega 的双荧光素酶报告分析系统。

2. 根据生产厂家的建议用光度计测量荧光。应根据经验确定光度计的线性范围。对于自动光度计，可以将与自动注射器相连的试管置于分析试剂中。

疑难解答

可能原因：荧光信号超过光度计的线性范围。

解决方案：用裂解缓冲液稀释样品。

方案 4.5 氯霉素乙酰转移酶分析

时线与安排

这里所描述的基本 CAT 分析方案是 Gorman 等（1982）方案的修改版。该方案监测通过 *E. coli* CAT 酶的［^{14}C］氯霉素乙酰化，其中乙酰化采用薄层层析（TLC）检测。如专题 4.3 中所描述的，已经发展了监测 CAT 活性的几种其他方法。特别是 Promega 销售的 CAT 酶分析系统（E1000），该系统可通过液体闪烁计数或 TLC 监测 CAT 活性。有关 TLC 方法和其他方法的其他信息，见 Sambrook 和 Russell（2001，pp. 17.33-17.41）、Ausubel 等（1994，单元 9），以及出售 CAT 分析试剂及其试剂盒的特定公司的相关资料。

TLC 方法以转染 CAT 报告质粒的细胞的裂解开始，通常通过冻融完成裂解，但是也可以用基于去污剂的缓冲液［如 Promega 的报告基因裂解缓冲液（E3971）］完成裂解。通过离心去除细胞碎片后，将裂解物和含有氯霉素和乙酰辅酶 A 的溶液一起孵育。通过使用乙酸乙酯的有机抽提，将乙酰化的氯霉素产物和未经修饰的反应物从水溶液中分离出来，然后通过冻干法去除乙酸乙酯，将氯霉素衍生物重悬于小体积乙酸乙酯中，并将其点到 TLC 板上。层析后，TLC 板在 X 光片上曝光，或采用磷屏成像仪分析。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

乙酰辅酶 A（5 mg/ml）

细胞，用目的质粒和 CAT 报告基因转染
氯霉素（^{14}C-标记的；54 mCi/mmol，25 μCi/ml）〈!〉
氯仿：甲醇（19：1）〈!〉
干冰〈!〉
乙醇〈!〉
乙酸乙酯〈!〉
磷酸缓冲盐（PBS）

试剂	10×PBS 原液（g/L）	1×PBS 工作液（mmol/L）
KCl〈!〉	2.0	2.7
KH_2PO_4〈!〉	2.0	1.4
$Na_2HPO_4 \cdot 7H_2O$〈!〉	11.5	4.3
NaCl	80.0	137.0

调节工作液的 pH 至 7.3。
蛋白质分析试剂（如 Bradford）
闪烁液（可选）
Tris-HCl（0.25 mol/L；pH 7.8 和 8.0）

设备

吹风机（可选，见步骤 2：6）
离心机（临床）（可选）
容器（玻璃）
滤纸（Whatman 3 MM）（可选）
通风橱
加热块，设置在 65℃
培养箱，设置在 37℃
冻干器（加热；如 Speed Vac 浓缩器）
微量离心机
磷屏成像仪（可选）
自动移液器
保鲜膜
剃刀刀片（可选）
闪烁计数器（可选）
闪烁瓶（可选）
TLC 板（可弯曲，20 cm × 20 cm，硅胶包被；Mallinckrodt Baker，4462-04）
试管（锥形，15 ml）（可选）
试管（离心，1.5 ml）

水浴，设置在 37℃

X 光片（可选）

程序

步骤 1：裂解细胞

1. 去除转染细胞的生长培养基。

2. 收集并裂解细胞。

对于贴壁细胞

a. 在培养皿中直接用 PBS 洗涤细胞。

b. 将细胞从培养皿上刮下，用自动移液器将细胞转移至装有小体积 PBS 的离心管中。

对于非贴壁细胞

a. 将非贴壁的细胞转移到 15 ml 管中，1500 r/min 离心 5 min，用 PBS 洗涤。

b. 使细胞重悬于小体积 PBS 中，将其转移到离心管中。

3. 在离心机中低速（8000 r/min）离心 1 min，沉淀细胞。弃上清液，并用 1 ml PBS 洗涤。如同上文沉淀细胞，并弃上清液。

4. 将细胞沉淀悬浮于 2 倍体积的 0.25 mol/L Tris-HCl（pH 8.0）中，涡旋。

5. 在干冰-乙醇浴中冰冻溶液，在 37℃水浴中融化，重复冻融循环三次，确保细胞完全裂解。

注意：也可以用含有去污剂的缓冲液裂解细胞。Promega 的报告基因裂解缓冲液（见方案 4.4）是为萤光素酶、CAT 和 β-半乳糖苷酶分析设计的。

6. 将悬液于 65℃加热 10 min。

注意：如果共转染编码 β-半乳糖苷酶的质粒用于归一化，那么省略此处理，因为热处理会使 β-半乳糖苷酶失活。或者，在热处理之前取出一份抽提物用于 β-半乳糖苷酶分析。

7. 在冰上冷却样品。在微量离心机中以 14 000 r/min 离心 5 min，沉淀细胞碎片。将上清液（含有可溶性 CAT 蛋白）转移至新管中，这时可在－70℃冷冻上清液。

8. 进行 Bradford 分析，确定各裂解物的总蛋白浓度。对于 CAT 分析，使用 75 μg 各裂解物。

步骤 2：CAT 分析

1. 对于每个反应，加入：

乙酰辅酶 A（5 mg/ml）	5 μl
[^{14}C] 氯霉素	2 μl
细胞裂解物	75 μg
0.25 mol/L Tris-HCl（pH 7.8）	至 125 μl

2. 37℃孵育 4 h。

注意：可根据信号强度改变反应时间。对于裂解物中 CAT 酶含量非常少的实验，反应可过夜进行。对于裂解物中 CAT 酶含量丰富的实验，反应可在 1 h 后终止。

3. 向每个反应加入 500 μl 乙酸乙酯，至少涡旋 10 s。

4. 室温下，在离心机中，14 000 r/min 离心 15 s，分相。

5. 转移上层（有机相）至新管中，并冻干（加热），沉淀可－70℃保存。

6. 将沉淀重悬于 15 μl 乙酸乙酯中。在硅胶 TLC 板上点样，斑点距板底部边缘 1 cm。缓慢点样，且不要将 15 μl 一次全部点上。斑点溶剂挥发后，再点剩余的样品。

注意：点样期间用吹风机可以加速溶剂的挥发。

7. 将板置于含有 19∶1 的氯仿∶甲醇溶液的封闭玻璃容器中。

注意：容器可内衬与接触洗脱液的 Whatman 3 MM 滤纸，以防止层析期间溶剂的挥发。

8. 溶剂前沿到达 TLC 板的 2/3 处时停止层析。将平板从容器中取出，在通风橱中面朝上风干，并包裹在塑料包装膜中。

9. 将平板在 X 光片上曝光或置于磷屏成像仪上过夜。

注意：有 4 个明显位于点样原点上方的斑点（图 4.6）。最低的斑点是非乙酰化的氯霉素，其次的两个斑点是单乙酰化产物，最上面的斑点是双乙酰化产物，该产物只有在存在大量酶活性时（而且当该分析超出线性范围时）才能观察到。

10. 在没有磷屏成像仪的情况下定量 CAT 活性：

a. 将 TLC 板和 X 光片对齐。

b. 用剃刀刀片切下斑点。

c. 将切下的斑点转移至闪烁瓶，加入闪烁液。

d. 用闪烁计数器计数。

注意：数据通常表示为百分比形式，相当于乙酰化斑点数除以总斑点数（乙酰化斑点加非乙酰化斑点）。

方案 4.6　β-半乳糖苷酶分析

时线与安排

虽然编码 β-半乳糖苷酶的 *E. coli lacZ* 基因像 CAT 和萤光素酶基因一样，可用作标准报告基因，以用于监测瞬时或稳定转染分析中启动子或增强子的强度，但是它主要用作瞬时转染实验期间的内对照（见专题 4.3 和本章中的内对照部分）。这里所描述的基本比色分析是用于定量 β-半乳糖苷酶活性的最简单、最廉价的分析。当 β-半乳糖苷酶报告质粒作为内对照（*lacZ* 基因处于强病毒启动子/增强子的控制之下）时，该方法经济又方便，但它不及萤光素酶或 CAT 分析灵敏，因此该方法对于弱活性或中等活性的启动子的监测没用。尽管如此，使用可商业性获得的化学发光分析和 ELISA，可以大大提高 β-半乳糖苷酶报告基因的灵敏性（见专题 4.4 和 Ausubel et al. 1994，单元 9）。

这里所描述的该分析的第一步是，裂解转染 β-半乳糖苷酶报告质粒的细胞。当用作

内对照时，细胞将已经转染了对照质粒和实验质粒，其中对照质粒含有融合于 *E. coli lacZ* 基因且存在遍在活性的病毒启动子，而实验质粒含有受目的启动子或增强子控制的另一个报告基因（如萤光素酶或 CAT）。可以用方案 4.5（及下文）所描述的冻融方法或方案 4.4 中所描述的去污剂方法裂解细胞。在确定了抽提物中总蛋白浓度（通常通过Bradford分析）后，将一份抽提物和反应底物邻硝基苯-β-D-吡喃半乳糖苷（ONPG）在含有磷酸钠和氯化镁的缓冲液中混合。当看到黄色产物时，用碳酸钠终止反应，并在波长 410 nm处确定样品的光密度。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

细胞，用目的质粒和 *lacZ* 报告基因转染的

干冰〈!〉

乙醇〈!〉

Mg^{2+}缓冲液

14 mol/L β-巯基乙醇〈!〉	350 μl
H_2O	550 μl
1 mol/L $MgCl_2$〈!〉	100 μl

邻硝基苯-β-D-吡喃半乳糖苷（ONPG），4.0 mg/ml，在 0.1 mol/L 磷酸缓冲液（pH7.5）中

磷酸缓冲液原液（0.8 mol/L）

试剂	g/L
Na_2HPO_4	95.44
$NaH_2PO_4 \cdot H_2O$	17.64

使用前按 1∶8 稀释。

磷酸缓冲盐（PBS）

试剂	10×PBS 原液/(g/L)	1×PBS 工作液/(mmol/L)
KCl〈!〉	2.0	2.7
KH_2PO_4〈!〉	2.0	1.4
$Na_2HPO_4 \cdot 7H_2O$〈!〉	11.5	4.3
NaCl	80.0	137.0

调节工作液的 pH 至 7.3。

蛋白分析试剂（如 Bradford）

Tris-HCl（0.25 mol/L；pH 8.0）

设备

离心机（临床）（可选）

孵育器，设置在 37℃
微量离心机
自动移液器
分光光度计
试管（锥形，15 ml）（可选）
试管（离心，1.5 ml）
水浴，设置在 37℃

程序

步骤 1：配制缓冲液

配制 Mg^{2+} 缓冲液、1×ONPG，以及 0.8 mol/L 磷酸缓冲盐（按 1∶8 稀释，配制 ONPG）。

步骤 2：裂解细胞

按照方案 4.5 的步骤 1～8 收集转染细胞，通过冻融法裂解细胞，并确定蛋白质浓度。

步骤 3：β-半乳糖苷酶分析

1. 为每个细胞裂解物建立反应体系（高度推荐重复样品）。

Mg^{2+} 缓冲液	3 μl
1×ONPG	66 μl
细胞裂解物	25 μg

加 0.1 mol/L 磷酸缓冲液（pH 7.5）至总体积 300 μl。

注意：由于许多哺乳动物细胞具有内源 β-半乳糖苷酶活性，所以包含阴性对照样品（来自于非转染细胞的裂解物）至关重要。

2. 37℃孵育，直至呈现淡黄色，至少在预期含有最大活性的管中出现淡黄色。孵育时间通常为 0.5～2 h。

3. 读取 420 nm 处的吸光度。

参考文献

Agalioti T., Lomvardas S., Parekh B., Yie J., Maniatis T., and Thanos D. 2000. Ordered recruitment of chromatin modifying and general transcription factors to the IFN-β promoter. *Cell* 103：667-678.

Annweiler A., Muller U., and Wirth T. 1992. Functional analysis of defined mutations in the immunoglobulin heavy-chain enhancer in transgenic mice. *Nucleic Acids Res*. 20：1503-1509.

Ansel K. M., Greenwald R. J., Agarwal S., Bassing C. H., Monticelli S., Interlandi J., Djuretic I. M., Lee D. U., Sharpe A. H., Alt F. W., and Rao A. 2004. Deletion of a conserved *Il4* silencer impairs T helper type1-mediated immunity. *Nat. Immunol*. 5：1251-1259.

Ausubel F. M., Brent R. E., Kingston E., Moore D. D., Seidman J. G., Smith J. A., and Struhl K. 1994. *Current protocols in molecular biology*. John Wiley and Sons, New York.

Banerji J., Olson J., and Schaffner W. 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* 33: 729-740.

Barton M. C., Madani N., and Emerson B. M. 1993. The erythroid protein GATA-1 functions with a stage-specific factor to activate transcription of chromatin-assembled β-globin genes. Genes *Dev.* 7: 1796-1809.

Barton M. C., Madani N., and Emerson B. M. 1997. Distal enhancer regulation by promoter derepression in topologically constrained DNA in vitro. *Proc. Natl. Acad. Sci.* 94: 7257-7262.

Bartsch J., Truss M., Bode J., and Beato M. 1996. Moderate increase in histone acetylation activates the mouse mammary tumor virus promoter and remodels its nucleosome structure. *Proc. Natl. Acad. Sci.* 93: 10741-10746.

Beard C., Hochedlinger K., Plath K., Wutz A., and Jaenisch R. 2006. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis* 44: 23-28.

Beato M. 1996. Chromatin structure and the regulation of gene expression: Remodeling at the MMTV promoter. *J. Mol. Med.* 74: 711-724.

Blochlinger K. and Diggelmann H. 1984. Hygromycin B phosphotransferase as a selectable marker for DNA transfer experiments with higher eucaryotic cells. *Mol. Cell. Biol.* 4: 2929-2931.

Blom van Assendelft G., Hanscombe O., Grosveld F., and Greaves D. R. 1989. The β-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. *Cell* 56: 969-977.

Bories J. C., Demengeot J., Davidson L., and Alt F. W. 1996. Gene-targeted deletion and replacement mutations of the T-cell receptor β-chain enhancer: The role of enhancer elements in controlling V (D) J recombination accessibility. *Proc. Natl. Acad. Sci.* 93: 7871-7876.

Boussif O., Lezoualc'h F., Zanta M. A., Mergny M. D., Scherman D., Demeneix B., and Behr J. P. 1995. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: Polyethylenimine. *Proc. Natl. Acad. Sci.* 92: 7297-7301.

Capecchi M. R. 1989. Altering the genome by homologous recombination. *Science* 244: 1288-1292.

Carey M., Leatherwood J., and Ptashne M. 1990. A potent GAL4 derivative activates transcription at a distance in vitro. *Science* 247: 710-712.

Chalfie M., Tu Y., Euskirchen G., Ward W. W., and Prasher D. C. 1994. Green fluorescent protein as a marker for gene expression. *Science* 263: 802-805.

Chaney W. G., Howard D. R., Pollard J. W., Sallustio S., and Stanley P. 1986. High-frequency transfection of CHO cells using polybrene. *Somat. Cell Mol.* Genet. 12: 237-244.

Charnay P., Mellon P., and Maniatis T. 1985. Linker scanning mutagenesis of the 5'-flanking region of the mouse β-major-globin gene: Sequence requirements for transcription in erythroid and nonerythroid cells. *Mol. Cell. Biol.* 5: 1498-1511.

Chittenden T., Lupton S., and Levine A. J. 1989. Functional limits of *oriP*, the Epstein-Barr virus plasmid origin of replication. *J. Virol.* 63: 3016-3025.

Chung J. H., Bell A. C., and Felsenfeld G. 1997. Characterization of the chicken β-globin insulator. *Proc. Natl Acad. Sci.* 94: 575-580.

de Wet J. R., Wood K. V., DeLuca M., Helinski D. R., and Subramani S. 1987. Firefly luciferase gene: Structure and expression in mammalian cells. *Mol. Cell. Biol.* 7: 725-737.

Emanuel J. R., Schulz X. M., and Zhou E. A. 1988. Expression an ouabain-resistant Na, K-ATPase in CV-1 cells after transfection with a cDNA encoding the rat Na, K-ATPase α1 subunit. *J. Biol. Chem.* 263: 7726-7733.

Ernst P. and Smale S. T. 1995. Combinatorial regulation of transcription Ⅱ: The immunoglobulin μ heavy chain gene. *Immunity* 2: 427-438.

Farr A. and Roman A. 1992. A pitfall of using a second plasmid to determine transfection efficiency. *Nucleic Acids Res.* 20: 920.

Felgner P. L., Gadek T. R., Holm M., Roman R., Chan H. W., Wenz M., Northrop J. P., Ringold G. M., and Danielsen M. 1987. Lipofection: A highly efficient lipid-mediated DNA transfection procedure. *Proc. Natl. Acad. Sci.* 84: 7413-7417.

Flanagan W. M. and Crabtree G. R. 1992. In vitro transcription faithfully reflecting T-cell activation requirements. *J. Biol. Chem.* 267: 399-406.

Forrester W. C., Novak U., Gelinas R., and Groudine M. 1989. Molecular analysis of the human β-globin locus activation region. *Proc. Natl. Acad. Sci.* 86: 5439-5443.

Fromm M. and Berg P. 1982. Deletion mapping of DNA regions required for SV40 early region promoter function in vivo. *J. Mol. Appl. Genet.* 1: 457-481.

Garraway I. P., Semple K., and Smale S. T. 1996. Transcription of the lymphocyte-specific terminal deoxynucleotidyltransferase gene requires a specific core promoter structure. *Proc. Natl. Acad. Sci.* 93: 4336-4341.

Gillies S. D., Morrison S. L., Oi V. T., and Tonegawa S. 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* 33: 717-728.

Gluzman Y. 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell 23: 175-182.

Goodbourn S., Zinn K., and Maniatis T. 1985. Human β-interferon gene expression is regulated by an inducible enhancer element. *Cell* 41: 509-520.

Gorman G. M., Moffat L. F., and Howard B. H. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* 2: 1044-1051.

Graham F. L. and van der Eb A. J. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* 52: 456-467.

Grosschedl R. and Baltimore D. 1985. Cell-type specificity of immunoglobulin gene expression is regulated by at least three DNA sequence elements. *Cell* 41: 885-897.

Gu H., Zou Y. R., and Rajewsky K. 1993. Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell* 73: 1155-1164.

Hall C. V., Jacob P. E., Ringold G. M., and Lee F. 1983. Expression and regulation of *Escherichia coli* lacZ gene fusions in mammalian cells. *J. Mol. Appl. Genet.* 2: 101-109.

Hartman S. C. and Mulligan R. C. 1988. Two dominant-acting selectable markers for gene transfer studies in mammalian cells. *Proc. Natl. Acad. Sci.* 85: 8047-8051.

Hsu L. Y., Lauring J. Liang H. E., Greenbaum S., Cado D., Zhuang Y., and Schlissel M. S. 2003. A conserved transcriptional enhancer regulates RAG gene expression in developing B cells. *Immunity* 19: 105-117.

Izumi M., Miyazawa H., Kamakura T., Yamaguchi I., Endo T., and Hanaoka F. 1991. Blasticidin S-resistance gene (bsr): A novel selectable marker for mammalian cells. *Exp. Cell Res.* 197: 229-233.

Janeway C. A. and Travers p. 1996. *Immunobiology: The immune system in health and disease*, 2nd ed. Current Biology, London and San Francisco; Garland, New York.

Jenuwein T. and Grosschedl R. 1991. Complex pattern of immunoglobulin p gene expression in normal and transgenic mice: Nonoverlapping regulatory sequences govern distinct tissue specificities. *Genes Dev.* 5: 932-943.

Kadonaga J. T. 1991. Purification of sequence-specific binding proteins by DNA affinity chromatography. *Methods Enzymol.* 208: 10-23.

Kraus W. L. and Kadonaga J. T. 1998. p300 and estrogen receptor cooperatively activate transcription via differential enhancement of initiation and reinitiation. *Genes Dev.* 12: 331-342.

Kushner P. J., Levinson B. B., and Goodman H. M. 1982. A plasmid that replicates in both mouse and *E. coli* cells. I. *Mol. Appl. Genet.* 1: 527-538.

Laybourn P. J. and Kadonaga J. T. 1992. Threshold phenomena and long-distance activation of transcription by RNA polymerase II. *Science* 257: 1682-1685.

Lee G. R., Spilianakis C. G., and Flavell R. A. 2005. Hypersensitive site 7 of the TH2 locus control region is essen-

tial for expressing TH2 cytokine genes and for long-range intrachromosomal interactions. *Nat. Immunol.* 6: 10-12.

Lenardo M., Pierce J. W., and Baltimore D. 1987. Protein-binding sites in Ig gene enhancers determine transcriptional activity and inducibility. *Science* 36: 1573-1577.

Leung T. H., Hoffmann A., and Baltimore D. 2004. One nucleotide in a κB site can determine cofactor specificity for NF-κB dimers. *Cell* 118: 453-464.

Liston D. R. and Johnson P. J. 1999. Analysis of a ubiquitous promoter element in a primitive eukaryote: Early evolution of the initiator element. *Mol. Cell. Biol.* 19: 2380-2388.

Liu H., Kang H., Liu R., Chen X., and Zhao K. 2002. Maximal induction of a subset of interferon target genes requires the chromatinremodeling activity of the BAF complex. *Mol. Cell. Biol.* 22: 6471-6479.

Liu J., O'Kane D. J., and Escher A. 1997. Secretion of functional *Renilla reniformis* luciferase by mammalian cells. *Gene* 203: 141-148.

Lo K., Landau N. R., and Smale S. T. 1991. LyF-1, a transcriptional regulator that interacts with a novel class of promoters for lymphocyte-specific genes. *Mol. Cell. Biol.* 11: 5229-5243.

Lomvardas S. and Thanos D. 2002. Modifying gene expression programs by altering core promoter chromatin architecture. *Cell* 110: 261-271.

Luthman H. and Magnusson G. 1983. High efficiency polyoma DNA transfection of chloroquine-treated cells. *Nucleic Acids Res.* 11: 1295-1308.

Malone C. S., Omori S. A., and Wall R. 1997. Silencer elements controlling the B29 (Igβ) promoter are neither promoter-nor cell-type-specific. *Proc. Natl. Acad. Sci.* 94: 12314-12319.

Mannino R. J. and Gould-Fogerite S. 1988. Liposome mediated gene transfer. *Biotechniques* 6: 682-690.

Mantei N., Boll W., and Weissmann C. 1979. Rabbit β-globin mRNA production in mouse L cells transformed with cloned rabbitβ-globin chromosomal DNA. *Nature* 281: 40-46.

Matthews J. C., Hon K., and Cormier M. J. 1977. Purification and properties of *Renilla reniformis* luciferase. *Biochemistry* 16: 85-91.

McCutchan J. H. and Pagano J. S. 1968. Enhancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethylaminoethyl dextran. *J. Natl. Cancer Inst.* 41: 351-357.

Mercola M., Goverman J., Mirell C., and Calame K. 1985. Immunoglobulin heavy-chain enhancer requires one or more tissue-specific factors. Science 227: 266-270.

Mohrs M., Blankespoor C. M., Wang Z. E., Loots G. G., Afzal V., Hadeiba H., Shinkai K., Rubin E. M., and Locksley R. M. 2001. Deletion of a coordinate regulator of type 2 cytokine expression in mice. *Nat. Immunol.* 2: 842-847.

Muljo S. A. and Schlissel M. S. 2002. The variable, C_H1, C_H2, and C_H3 domains of Ig heavy chain are dispensable for pre-BCR function in transgenic mice. *Int. Immunol.* 14: 577-584.

Mulligan R. C. and Berg P. 1981. Selection for animal cells that express the *Escherichia coli* gene coding for xanthine-guanine phosphoribosyltransferase. *Proc. Natl. Acad.* Sci. 78: 2072-2076.

Mulsant P., Gatignol A., Dalens M., and Tiraby G. 1988. Phleomycin resistance as a dominant selectable marker in CHO cells. *Somatic Cell Mol. Genet.* 14: 243-252.

Neumann E., Schaefer-Ridder M., Wang Y., and Hofschneider P. H. 1982. Gene transfer into mouse lyoma cells by electroporation in high electric fields. *EMBO J.* 1: 841-845.

Nguyen V. T., Morange M., and Bensaude O. 1988. Firefly luciferase luminescence assays using scintillation counters for quantitation in transfected mammalian cells. *Anal. Biochem.* 171: 404-408.

Ohe Y., Zhao D., Saijo N., and Podack E. R. 1995. Construction of a novel bovine papillomavirus vector without detectable transforming activity suitable for gene transfer. *Hum. Gene Ther.* 6: 325-333.

Parker B. A. and Stark G. R. 1979. Regulation of simian virus 40 transcription: Sensitive analysis of the RNA spe-

cies present early in infections by virus or viral DNA. *J. Virol.* 31: 360-369.

Pazin M. J., Sheridan P. L., Cannon K., Cao Z., Keck J. G., Kadonaga J. T., and Jones K. A. 1996. NF-κB-mediated chromatin reconfiguration and transcriptional activation of the HIV-1 enhancer in Vitro. *Genes Dev.* 10: 37-49.

Pierani A., Heguy A., Fujii H., and Roeder R. G. 1990. Activation of octamer-containing promoters by either octamer binding transcription factor 1 (OTF-1) or OTF-2 and requirement of an additional B-cell-specific component for optimal transcription of immunoglobulin promoters. *Mol. Cell. Biol.* 10: 6204-6215.

Plevy S. E., Gemberling J. H. M., Hsu. S., Dorner A. J., and Smale S. T. 1997. Multiple control elements mediate activation of the murine and human interleukin 12 p40 promoters: Evidence of functional synergy between C/EBP and Rd proteins. *Mol. Cell. Biol.* 17: 4572-4588.

Potter H. 1988. Electroporation in biology: Methods, applications, and instrumentation. *Anal. Biochem.* 174: 361-373.

Prasher D. C., Eckenrode V. K., Ward W. W., Prendergast F. G., and Cormier M. J. 1992. Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* 111: 229-233.

Promega Corporation. 1996. *Protocols and applications guide*, 3rd ed. Promega, Madison, Wisconsin.

Robbins P. D., Rio D. C., and Botchan M. R. 1986. *trans* Activation of the simian virus 40 enhancer. *Mol. Cell. Biol.* 6: 1283-1285.

Sambrook J. and Russell D. W. 2001. *Molecular cloning: A laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schmidt E. V., Christoph G., Zeller R., and Leder P. 1990. The cytomegalovirus enhancer: A pan-active control element in transgenic mice. *Mol. Cell. Biol.* 10: 4406-4411.

Simonsen C. C. and Levinson A. D. 1983. Isolation and expression of an altered mouse dihydrofolate reductase cDNA. *Proc. Natl. Acad. Sci.* 80: 2495-2499.

Smith C. L. and Hager G. L. 1997. Transcriptional regulation of mammalian genes in vivo: A tale of two templates. *J. Biol. Chem.* 272: 27493-27496.

Solymar D. C., Agarwal S., Bassing C. H., Alt F. W., and Rao A. 2002. A 3′ enhancer in the IL-4 gene regulates cytokine production by Th2 cells and mast cells. *Immunity* 17: 41-50.

Southern P. J. and Berg P. 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *J. Mol. Appl. Genet.* 1: 327-341.

Spector D. L., Goldman R. D., and Leinwand L. A. 1998. Light microscopy and cell structure. In *Cells: A laboratory manual*, vol. 2., pp. 86. 1-86. 6. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Sternberg N. and Hamilton D. 1981. Bacteriophage P1 site-specific recombination: Recombination between *loxP* sites. *J. Mol. Biol.* 150: 467-486.

Stirland J. A., Seymour Z. C., Windeatt S., Norris A. J., Stanley P., Castro M. G., Loudon A. S., White M. R., and Davis J. R. 2003. Realtime imaging of gene promoter activity using an adenoviral reporter construct demonstrates transcriptional dynamics in normal anterior pituitary cells. *J. Endocrinol.* 178: 61-69.

Talbot D. and Grosveld F. 1991. The 5′ HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. *EMBO J.* 10: 1391-1398.

Taniuchi I., Sunshine M. J., Festenstein R., and Littman D. R. 2002. Evidence for distinct CD4 silencer functions at different stages of thymocyte differentiation. *Mol. Cell* 10: 1083-1096.

Treisman R. 1986. Identification of a protein-binding site that mediates transcriptional response of the *c-fos* gene to serum factors. *Cell* 46: 567-574.

Treisman R., Green M. R., and Maniatis T. 1983. *cis* and *trans* activation of globin gene transcription in transient assays. *Proc. Natl. Acad. Sci.* 83: 7428-7432.

Trinh L. A., Ferrini R., Cobb B. S., Weinmann A. S., Hahm K., Ernst P., Garraway I. P., Merkenschlager M., and Smale S. T. 2001. Down-regulation of TDT transcription in CD4$^+$CD8$^+$ thymocytes by Ikaros proteins in direct competition with an Ets activator. *Genes Dev*. 15: 1817-1832.

van der Weyden L., Adams D. J., and Bradley A. 2002. Tools for targeted manipulation of the mouse genome. *Physiol. Genomics* 11: 133-164.

Vara J. A., Portela A., Ortin J., and Jimenez A. 1986. Expression in mammalian cells of a gene from *Streptomyces alboniger* conferring puromycin resistance. *Nucleic Acids Res*. 14: 4617-4624.

Wigler M., Silverstein S., Lee L. S., Pellicer A., Cheng Y. C., and Axel R. 1977. Transfer of purified herpes virus thymidine kinase gent to cultured mouse cells. *Cell* 11: 223-232.

Wirth T., Staudt L., and Baltimore D. 1987. An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoidspecific promoter activity. *Nature* 329: 174-178.

Wright S., deBoer E., Grosveld F. G., and Flavell R. A. 1983. Regulated expression of the human β-globin gene family in murine erythroleukaemia cells. *Nature* 305: 333-336.

Wu H., Liu X., and Jaenisch R. 1994. Double replacement: Strategy for efficient introduction of subtle mutations into the murine *Colla*-1-gene by homologous recombination in embryonic stem cells. *Proc. Natl. Acad. Sci.* 91: 2819-2823.

Xu Y., Davidson L., Alt F. W., and Baltimore D. 1996. Deletion of the Igκ light chain intronic enhancer/matrix attachment region impairs but does not abolish VκJκ rearrangement. *Immunity* 4: 377-385.

Yates J. L., Warren N., and Sugden B. 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. *Nature* 313: 812-815.

Young D. C., Kingsley S. D., Ryan K. A., and Dutko F. J. 1993. Selective inactivation of eukaryotic 3-galactosidase in assays for inhibitors of HIV-1 TAT using bacterial 3-galactosidase as a reporter enzyme. *Anal. Biochem.* 215: 24-30.

Zheng C., Hoque A. T., Braddon V. R., Baum B. J., and O'Connell B. C. 2001. Evaluation of salivary gland acinar and ductal cell-specific promoters in vivo with recombinant adenoviral vectors. *Hum. Gene Ther.* 12: 2215-2223.

Zhong X. -P. and Krangel M. S. 1997. An enhancer blocking element between α and δ gene segments within the T cell receptor α/δ locus. *Proc. Natl. Acad. Sci.* 94: 5219-5224.

Zhou L., Nazarian A. A., Xu J. Tantin D., Corcoran L. M., and Smale ST. 2007. An inducible enhancer required for *Il12b* promoter activity in an insulated chromatin environment. *Mol. Cell. Biol.* 27: 2698-2712.

Zhu H., Yang J., Murphy T. L., Ouyang W., Wagner F., Saparov A., Weaver C. T., and Murphy K. M. 2001. Unexpected characteristics of the IFN-γ reporters in nontransformed T cells. *J. Immunol.* 167: 855-865.

Zimmermann U. 1982. Electric field-mediated fusion and related electrical phenomena. *Biochim. Biophys. Acta* 694: 227-277.

Zinn K., Mellon P., Ptashne M., and Maniatis T. 1982. Regulated expression of an extrachromosomal human β-interferon gene in mouse cells. *Proc. Natl. Acad. Sci.* 79: 4897-4901.

网络资源

http://www.atcc.org/：美国模式培养物保藏所（American Type Culture Collection）主页，转化细胞系的商业来源。

http://ccr.coriell.org/Sections/Collections/NIGMSI? SsId=8：由科里尔医学研究所（Coriell Institute for Medical Research）运营的国家普通医学研究所（National Institute of General Medical Sciences）细胞库主页。

5　远程控制区的鉴定和分析

要点

- 远程控制区对于正确的基因调控来说尽管不是始终需要的，但也是经常需要的。
- 远程控制区难以发现。
- 可用系统发生比较、DNase I 高敏感性分析、依赖于染色质免疫沉淀的组蛋白修饰分析、核基质结合分析，以及功能性方法等鉴定远程控制区。
- 要确证假定控制区对于正确的基因调控是重要的，就需要功能分析。

引言和概述

正如第 2 章中所描述的，可以以启动子或远程控制区为重点开始转录调控的分析，这两种策略都可能洞悉目的基因转录调控的机制。启动子通常提供一个比较容易的出发点，因为它的鉴定仅依赖于转录起始位点的确定。远程控制区比较难以鉴定，但是许多研究者都将它们视为转录调控的主要决定因素。

目前的模型认为在启动子与转录因子和 RNA 聚合酶结合之前，必须改变染色质高级结构。远程控制区可能通过影响染色质结构参与调控基因座的可及性（第 1 章）。因此，远程控制区在调控等级中可能是启动子的上游。此外，正如第 1 章和下文（也见第 9 章）所述，诸如染色质构象捕获（3C）分析的体内分析已揭示出，基因的远程控制区相互之间及其启动子之间物理空间上靠得很近，这种靠近主要借助这些区域间形成环，即成环，据推测这些区域是转录活化所必需的（Dekker et al. 2002；Toihuis et al. 2002；Spilianakis and Flavell 2004；de Laat et al. 2008）。还发现已被称为活性染色体中心的含有启动子和增强子的核点，与相似或者差异调控的其他基因的控制区靠得很近（Spilianakis et al. 2005；de Laat et al. 2008）。因此，与其将远程控制区视为调控等级中启动子的上游，倒不如说几个控制区同时指挥每个基因的正确调控；同时，一个有意义的分析将要求采用评价启动子与远程控制区物理上和功能上的相互作用的分析方法，

进行启动子和所有远程控制区的同步表征。

已经鉴定了许多类型的远程控制区，包括增强子、沉默子、基因座控制区（LCR）、边界元件及基质附着区（MAR）（见第1章）。在后生动物中，除启动子外，每种蛋白质编码基因都可能含有一个或多个这些区域。然而，每种类型的远程区所调控的基因的比例尚未确定。而且，对于任何基因，有助于其正确调控的远程控制区的个数也尚未明确地确定。

正如第1章和下文中所述，研究远程控制区的一个挑战是它们的界定特征不精确。例如，有些LCR只在染色体环境中起作用，而其他一些在瞬时分析中的附加体（episome）上起作用。此外，有些边界元件不具有增强子功能，而其他一些边界元件则表现出绝缘子和增强子的双重活性。MAR表现出广泛的功能。负责这些可变特征的分子机制尚不清楚。尽管在理解由DNA结合蛋白引起的转录活化的基本机制及染色质结构在基因调控中的作用方面已获得很大进展，但对远程控制区调控转录的精确机制仍然知之甚少。这种复杂性和不确定性的一个可能原因是多个远程控制区共同作用去协调附近基因的正确调控。因此，虽然个别控制区的研究是必不可少的，但是当它们被独立研究时，难以全面获悉它们的精确功能。

像启动子分析一样（见第4章），单个远程控制区的分析严重依赖于功能分析方法的发展。在很多情况下，这些功能分析方法与用于研究和解析启动子的功能分析方法相似，如瞬时转染和稳定转染分析，但是通常也需要难度更大、耗时更长的分析方法，如转基因分析方法、通过同源重组已经消除内源性控制区的小鼠或细胞的分析。这些难度更大、耗时更长的分析方法的研究对于理解控制区在其自然基因组环境中如何促成基因调控是必要的。

启动子和远程控制区之间的另一个重要区别关系到用于它们的初始鉴定和定位的策略。要鉴定启动子，简单地说就是要定位转录起始位点，因为启动子和起始位点从定义上说是紧密相连的（见第1章）。用于鉴定远程控制区的方法则截然不同。一个强烈推荐的出发点是进行基因座的系统发生分析，以鉴定进化中保守的基因附近的非编码序列。已经出版了许多用于进行比较基因组学的非常优秀的逐步指导（Nardone et al. 2004；Bejerano et al. 2005；Karolchik et al. 2007，2008）。

鉴定和评估真核基因组内保守的非编码区的两个首选工具是UCSC（University of California，Santa Cruz）基因组浏览器（http://genome.ucsc.edu）（Hinrichs et al. 2006；Karolchik et al. 2007，2008）和Ensembl基因组浏览器（www.ensembl.org）。不幸的是，保守区的长度及预测存在功能性相关远程控制区的保守程度尚未确定。但是，正在出现各种各样的计算策略，以用于评估脊椎动物基因组非编码序列的保守性。图5.1显示了一个由UCSC基因组浏览器产生的系统发生图谱的例子，该图谱显示了小鼠促炎症反应细胞因子基因 *Il12b* 附近的保守元件（在第6章更深入讨论）。图5.1中的图谱也显示了小鼠 *Il12b* 基因座与部分其他脊椎动物物种中同源基因座之间的保守性。通过改变浏览器设置，可以显示更多或更少的物种。顶部显示了哺乳动物物种总体保守性。在基因组浏览器显示的底部，显示了序列模体；通过已测序的所有哺乳动物物种、所有真祖兽总目物种（灵长类动物、啮齿类动物和兔子）及所有脊椎动物物种的比

较，这些模体显示出特别强的概率对数得分。在浏览器网站（显示目的基因座后，点击比较基因组学部分中的“保守”或者“最保守”链接）或者 David Haussler 实验室的出版物中，可以发现关于 UCSC 基因组浏览器现有功能的详细描述（Bejerano et al. 2005；Karolchik et al. 2007，2008）。在核酸研究的“数据库专期（Database Issue）”上每年都公布浏览器功能的更新。

在图 5.1 中的 *Il12b* 基因座的系统发生图谱揭示了与在高等真核生物中大多数基因座中发现的类似的基因组特征。在蛋白质编码外显子处（顶部矩形）和启动子内（底部简图中接近位置 44214000 的实框），强保守性明显。此外，三个 lod 图（标注为哺乳动物、真祖兽总目和脊椎动物）中的每条垂直线都描绘了单个的 DNA 元件，这些元件在测序的哺乳动物、真祖兽总目或脊椎动物中特别保守。几个 lod 评分很高的成簇元件很明显位于转录起始点（底部简图中接近位置 44214000 的实框）上游约 10 kb 处，这一区域对应于 *Il12b* 增强子，并且最初鉴定为一个强 DNase I 高敏感性位点，但该位点只在已被刺激以诱导 *Il12b* 转录的细胞里才显示高敏感性（Zhou et al. 2007）。有趣的是，系统发生分析揭示了其他三个介于启动子和－10 kb 增强子之间的区域，该区域含有高 lod 得分的元件（底部简图的三个虚框）。位于转录起始位点上游约 1 kb 的具有高 lod 评分的成簇元件（位置 4413000 附近的三个虚框）表现出组成型核酸酶高敏感性（Weinmann et al. 1999）。但是，在其他两个区域（位置 44206000 和 44209000 附近第一和第二个虚框）没有观察到核酸酶高敏感性（Zhou et al. 2007）。在 *Il12b* 内含子

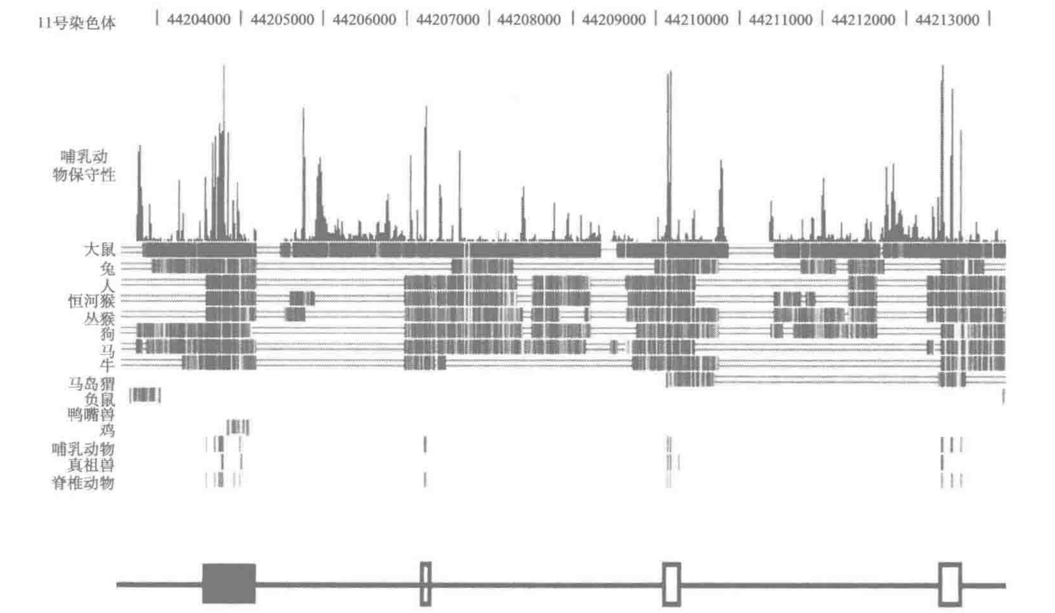


图 5.1　使用 UCSC 基因组浏览器（http://genome.ucsd.edu）对小鼠 *Il12b* 基因座进行系统发生分析（Kent et al. 2002）。顶部显示的是 11 号染色体上的核苷酸位置，核苷酸位置下方描述的是 *Il12b* 的外显子和内含子（矩形和阴影线）。接下来显示的是总体哺乳动物保守性（Mammal Cons.），这种保守性基于对多种哺乳动物物种的序列比对，序列比对采用的工具是 Siepel 等（2005）开发的 phastCons。保守性轨道下面显示的是小鼠序列相对于个体脊椎动物物种序列的保守性。

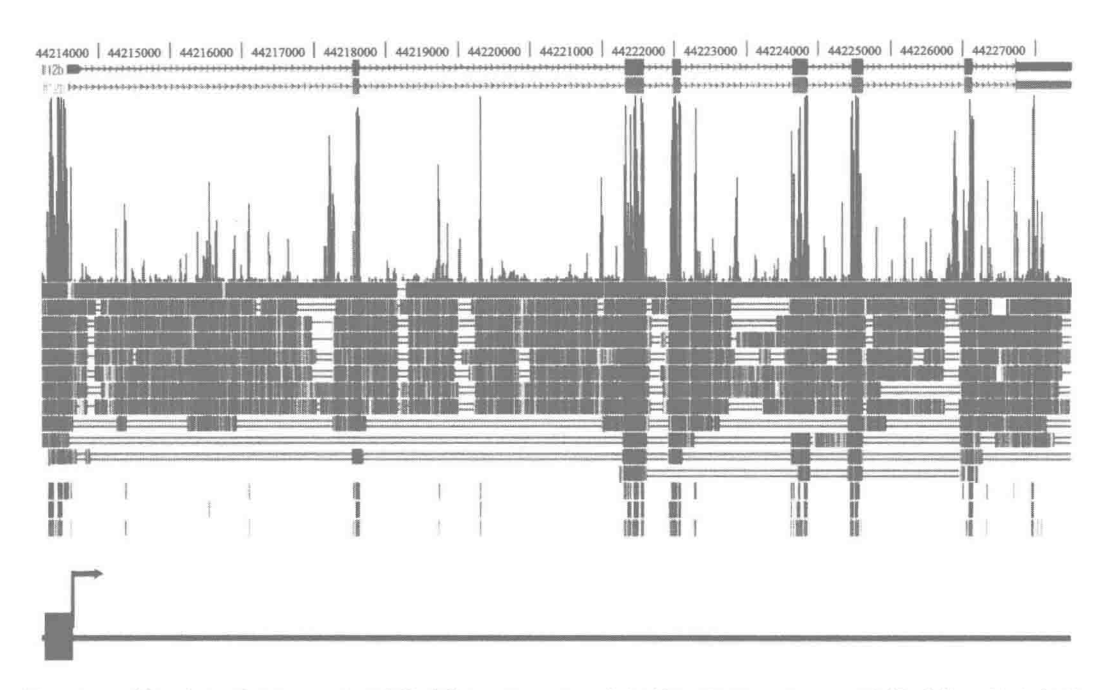


图 5.1　(续) 标记为 Mammal (哺乳动物)、Euarch (真祖兽) 和 Vertebrate (脊椎动物) 的底部轨道，以垂直线描绘了 DNA 模体，与数据库中不同物种的序列分别作比较，这些模体显示出高概率对数 (lod) 值；其中数据库来自数种哺乳动物物种、真祖兽总目物种 (灵长类动物、啮齿类动物和兔子) 或脊椎动物物种。预测这些模体最有可能显示功能重要性 (Bejerano et al. 2005；Siepel et al. 2005)。用“比较基因组 (comparative genomics)”标题下面的“最保守 (most conserved)”设置就可以显示这个底部轨迹。在基因组浏览图谱下面的简图显示 *Il12b* 转录起始位点 (箭头)，还有 *Il12b* 启动子和－10 kb 增强子 (两个实框)，以及在启动子和增强子之间显示高 lod 评分的其他三个区域 (虚框)。

(图 5.1) 和 *Il12b* 转录单位的下游或－10 kb 上游的其他基因组区域也能发现具有高 lod 评分的许多其他的元件 (未显示)。

Il12b 基因座的上述特征表明了系统发生分析作为研究远程控制区研究出发点的价值，但它们也揭示了这种方法的局限性。系统发生分析成功地揭示了－10 kb 处的保守区域，该区域现在已知作为一个可诱导增强子起作用 (Zhou et al. 2007)。此分析还揭示了－1 kb 处组成性可及 DNA 区域内的保守元件，该元件可能对 *Il12b* 转录是重要的 (Weinmann et al. 1999)。此外，系统发生分析揭示了 *Il12b* 基因座附近的其他几个元件，这些元件可能促成转录调控。

这种方法的局限性在于，有了如此大量的保守元件，仅从系统发生分析难以预测哪些元件可能是 *Il12b* 转录的关键调控者。在自然染色质位置中，有可能所有这些保守元件都是 *Il12b* 基因的正确调控所需要的，并且对于它们应该全部同等重视。另外，有可能其他方法，如 DNase I 高敏感性，将更成功地鉴定对正确调控最重要的保守区域。例如，DNase I 高敏感性研究揭示了－10 kb 区域只在细胞活化后表现出核酸酶高敏感性，提

示这一区域可能对可诱导转录特别重要，而这一提示在只用系统发生分析时是无法提供的。

尽管有这一缺陷，单个 DNA 元件或者扩展的非编码区域的高度进化保守性，可能为在转染或者转基因分析中检验该元件或区域的功能相关性或者通过同源重组干扰该区域提供充分的理由。另外，当用其他策略鉴定远程控制区时（下文描述），系统发生保守性可以确实地激励研究者特别密切关注那个区域。

核酸酶高敏感性分析是一种用于鉴定远程控制区的实验方法，该方法已被非常成功地应用了 20 多年（见专题 5.1 及下文）。此分析与邻近目的基因的 DNA 区域的鉴定有关，这些 DNA 区域对内切酶，主要是 DNase I 消化有高敏感性。核酸酶高敏感性表明该区域无核小体，或者该区域被组装成结构改变的（或重塑的）核小体，正如所预期的与一群序列特异性 DNA 结合蛋白结合的远程控制区。因为这一特性，已经证实核酸酶高敏感性是定位远程控制区的一种相对快速的方法。

常规 DNase I 高敏感性方法致力于单个基因座。但是，目前已经发展了数种方法，用于目的细胞类型或者细胞群体中 DNase I 高敏感性位点的全基因组鉴定，这些方法依赖于叠连矩阵、高通量测序，或者其他策略（Crawford et al. 2004，2006a，b；Dorschner et al. 2004；Sabo et al. 2004，2006；Follows et al. 2006；Boyle et al. 2008）。某些情况下，如果目的细胞类型已经用这些方法之一事先分析过，那么仅仅通过搜索全基因组高敏感位点数据库，就有可能鉴定基因座附近的高敏感位点。

专题 5.1　DNase I 高敏感性分析

Southern 印迹

要进行 DNase I 高敏感性分析（图 5.2），就需要表达目的基因的细胞。一般对于每个 DNase I 滴定点来说需要 1×10^7 个细胞。将生长旺盛的细胞冷却、漂洗，用非离子型去污剂 NP-40 裂解以制备细胞核，然后用不同浓度的 DNase I 处理各份分离的细胞核。此外，也要设置不同的时间点，通过加入 EDTA 和 SDS 终止核酸酶消化。然后加入蛋白酶 K 消化细胞核蛋白和染色体蛋白（这时，DNase I 消化的大概程度变得清晰可见：低浓度 DNase I 或无 DNase I 处理的溶液由于分子质量大的 DNA 分子的存在而变得黏稠，用相对高浓度 DNase I 处理的样品不黏稠）。蛋白酶 K 处理后，通过酚和氯仿抽提，以及 RNase A 消化来纯化 DNA 样品。最后用乙醇沉淀 DNA，并重悬于水中。该程序的一个挑战是从黏稠和不太黏稠的样品中制备 DNA 期间，难以保持一致性，这是因为黏稠的样品将更难抽提，并且沉淀后也难悬浮。

为了采用 Southern 印迹分析切割片段，接下来用选择好的限制酶（通常是识别6～8 bp 序列的酶）对各份纯化的 DNA 样品进行消化，这些酶能在目的区域中产生5～20 kb的片段。将酶切产生的片段在低浓度琼脂糖凝胶上电泳，使各片段分离，并转移到适当的膜上，用于 Southern 分析。可以依次利用不同的放射性标记标记探针探测单张膜。通过使用不同的探针，可以评定跨越基因座的不同限制性片段中高敏感位点的存在。在基因上游几十 kb 处或基因下游及内含子中都已鉴定到了功能性控制元件。详

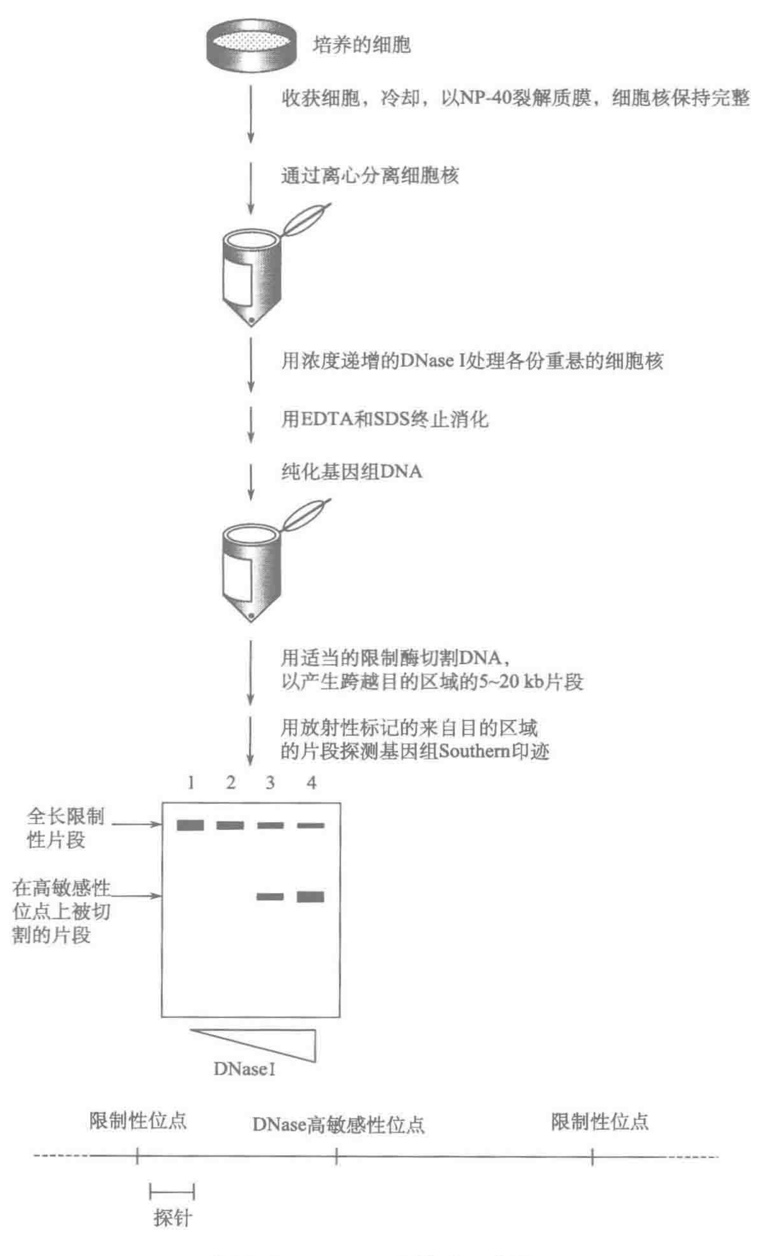


图 5.2 DNase I 高敏感性分析。

细的方案，见 Enver 等（1985）及正文中引用的主要研究论文，也可以按照方案 9.1（第 9 章）进行，只是用 DNase I 代替微球菌核酸酶。

体内微球菌核酸酶分析同样也可能为揭示关键调控区的位置提供信息，这种分析（方案 9.1）通过与 DNase I 高敏感性分析相同的方法进行。微球菌核酸酶具有不同于 DNase I 的特性，因为它优先在核小体连接区和无核小体的其他区域中引入双链 DNA

断裂。因此，它对于研究核小体结构和定位来说是一种很有价值的试剂（见第 9 章）。另外，一些重要的控制区对微球菌核酸酶和 DNase I 的切割都高度敏感。当进行核酸酶高敏感性分析时，试用这两种核酸酶是有益的，因为其中一种可能会比另一种产生更有说服力的结果。对于 DNase I 高敏感性研究，已经开发出基于聚合酶链反应（PCR）的技术替代 Southern 印迹方法。

全基因组方法

已经建立了多种全基因组方法，用于鉴定在目的细胞群中显示 DNase I 高敏感性的区域（Crawford et al. 2004，2006a，b；Dorschner et al. 2004；Sabo et al. 2004，2006；Follows et al. 2006；Boyle et al. 2008）。这些方法的其中一种是先如上所述将 DNA 消化，但是之后连接生物素标记的接头上（Crawford 2006a；Boyle et al. 2008）。原则上，接头将只能连接到 DNase I 消化的 DNA 末端，然后将 DNA 用超声波打碎成小片段，并将生物素化片段连接到链亲和素珠子上，这一步富集被 DNase I 切割的片段，去除缺少生物素接头和 DNase I 切割位点的超声波打碎的 DNA。然后将生物素化的 DNA 连接到第二组接头上，PCR 扩增，并用大规模平行 DNA 测序技术［如 Solexa（Illumina）或 454 Life Sciences（Roche）］测序或者与全基因组微阵列芯片杂交。

在一项人 $CD4^{+}$ T 细胞的全基因组分析中（Boyle et al. 2008），注释了 94 925 个 DNase I 高敏感位点。用于这项分析的方法涉及将生物素化的接头连接到在完整细胞核内以 DNase I 切割的基因组 DNA 上，最后同时用大规模平行测序和叠连阵列杂交分析（见图 5.1）。这项研究富有吸引力的特色是使用了两种分析方法并进行比较。有趣的是，在鉴定的 DNase I 高敏感位点中，只有 43%的位点与基因组的进化约束区（ECR）一致，这暗示在人 T 细胞中大部分重要调控区在进化中并不保守，或者大部分 DNA 区域由于尚未完全了解的原因对核酸酶裂解有高敏感性。

这项研究包括了高敏感性数据和最近的全基因组组蛋白修饰数据的有趣比较（Barski et al. 2007）。这些由 Boyle 等（2008）收集的高敏感性数据集可以很容易地在 UCSC 基因组浏览器（http://genome.ucsc.edu/cgi-bin/hgTrackUi？hgsid= 105645537&c=chr20&g=dukeDnaseCd4Sites 上 Duke DNase I 轨道）上访问。很可能在多个物种中，不久将获得几个其他细胞类型的综合 DNase I 高敏感性数据集。

作为鉴定远程控制区的另一种策略，染色质免疫沉淀（ChIP）可以用于扫描目的基因侧翼的基因组 DNA，以鉴定具有特定共价修饰的组蛋白富集的区域（如见 Roh et al. 2005；Szutorisz et al. 2005），这些区域可能对应于功能重要的远程控制区。现在已公认，诸如组蛋白 H3 和 H4 乙酰化及组蛋白 H3 赖氨酸 4 甲基化的组蛋白修饰，富集在活性基因或者准备激活基因的启动子和增强子处（Heintzman et al. 2007；Roh et al. 2007）（见第 1 章）。因此，如果为 ChIP 设计了 PCR 引物，并且这些引物用于监测目的基因侧翼以 100～500 bp 递增的基因组 DNA，那么 PCR 检测结果可能揭示与远程控制区对应的 DNA 区域。以类似的方式，如果认为特定转录因子调控目的基因，用于监测这些转录因子结合的 ChIP 分析，也可以用于扫描可能与该因子结合的远程控制区（如见 Grass et al. 2003）。

尽管用常规 ChIP 分析扫描目的基因座以发现与组蛋白修饰或转录因子结合的区域一般是适宜的，但是也可以从全基因组 ChIP 研究组装的数据库中获得有用的信息；在这些全基因组 ChIP 研究中，通过微阵列杂交（ChIP-chip）、基于系列扩增的全基因组定位技术（GMAT）或高通量平行测序（ChIP-Seq）来分析来自免疫沉淀染色质的 DNA（Rob et al. 2005，2006，2007；Barski et al. 2007；Heintzman et al. 2007；Roh and Zhao 2007；Schones and Zhao 2008）。如果对目的细胞类型已进行了全基因组研究，目前可用的数据库对于研究者来说是最有用的。但是，由于高通量平行测序的成本正在迅速降低，即使目标是仅仅分析单一的目的基因座，进行全基因组 ChIP-Seq 分析以检验以前未分析过的一个细胞类型中的组蛋白修饰，也是越来越可行。第 9 章中更详细地描述了 ChIP 分析，但是上面引用的论文提供了关于使用这种方法来鉴定潜在控制区的详细信息。

在缺乏之前核酸酶高敏感性或者 ChIP 研究的情况下，也可以通过功能分析，如转染分析或转基因小鼠分析，来鉴定一个基因周围的远程控制区。在一项对小鼠白细胞介素 4（*Il4*）基因座（下文讨论）的杰出分析中，用转基因小鼠分析鉴定出了一个关键的 LCR，而在先前的系统发生和 DNase I 高敏感性研究中则未鉴定出这一控制区（Lee et al. 2003；Smale 2003）。这个例子很好地显示出功能性方法和使用多种独立的实验策略来解决一个问题的价值。

除了以上策略外，也可以用基于 3C 的分析来扫描与目的基因的启动子或者已知的增强子紧密接近的区域（Dekker et al. 2002；Tolhius et al. 2002；Simonis et al. 2007）。关于 3C 分析的描述和相关分析，诸如 4C 和 5C，见第 9 章。最后，可以根据 MAR 与核基质的物理性结合来鉴定 MAR。有时，MAR 自身具有功能活性或者与重要的控制区有着密切的联系。虽然此时未强烈建议（见下文），但 MAR 的鉴定可能导致远程控制区的鉴定。

下面，我们聚焦于用于鉴定远程控制区的 DNase I 高敏感性分析和功能性策略，出于它们的历史重要性及当前它们的功能意义的不确定性，我们对 MAR 也做了简单的讨论。使用上述所列的任何一种策略可能鉴定到不同类型的控制区，我们对用于研究这些控制区的功能性方法做一讨论，以此结束本章。

实验策略

DNase I 高敏感性

DNase I 敏感性和高敏感性的基本原理

自 20 世纪 70 年代早期以来，就已经知道基因组 DNA 组装成核小体，核小体可进一步形成染色质高级结构（关于早期进展的综述，见 Elgin and Weintraub 1975）。从这一认识引发的一个问题是：活性基因上的染色质和非活性基因上的染色质是否相同。通过追问在分离的细胞核中活性基因和非活性基因相联系的染色质是否同样被胰 DNase I 切割，Weintraub 和 Groudine（1976）解决了这一问题。他们的结果揭示，与来自不表达珠蛋白的其他细胞类型的细胞核相比，从鸡红细胞分离的细胞核中的鸡珠蛋白基因编

码区，对 DNase I 消化更敏感。"DNase I 敏感性分析"（图 5.3）需要用浓度递增的 DNase I 处理分离的细胞核。在最初的研究中（Weintraub and Groudine 1976），利用溶液杂交方法确定了珠蛋白基因座和对照基因座的消化程度。

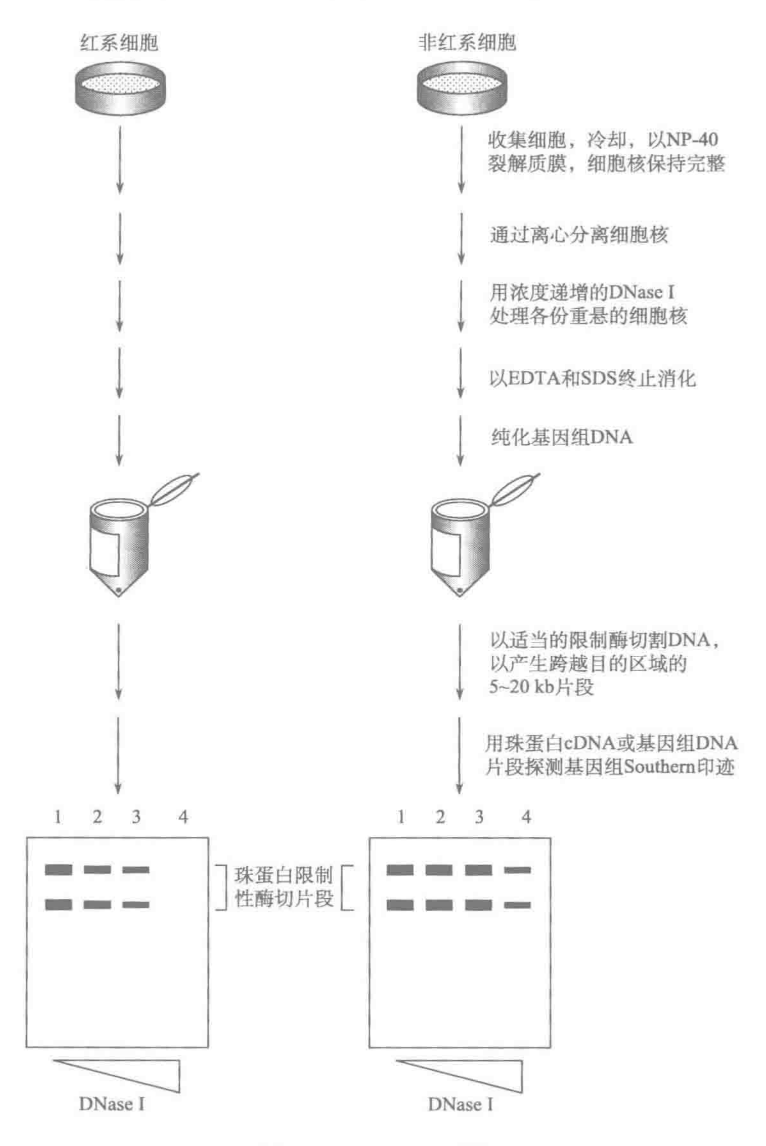


图 5.3 DNase 敏感性。

但是，溶液杂交不久被基因组 Southern 印迹所取代（Stalder et al. 1980）。对于这种分析方法，纯化经 DNase 消化的基因组 DNA，并用在珠蛋白基因座内产生已知片段大小的限制酶酶切片段，然后通过琼脂糖凝胶电泳分离 DNA 片段，并将其转移至硝酸纤维素膜上。通过与来源于珠蛋白基因或对照基因的放射性标记探针杂交，检测特定的限制性片段。在来自红细胞核的样品中，DNase I 有效地消化了珠蛋白基因座中的

DNA 序列（图 5.3，左），从基因组 Southern 印迹中相应限制性片段的缺失可明显地看出这一点。相反，在不表达珠蛋白的细胞类型中，珠蛋白 DNA 序列对 DNase I 有抗性（图 5.3，右）。而且在红细胞样品中，无转录活性的基因对 DNase I 也有抗性。从这些数据，Weintraub 和 Groudine 推断，当基因活跃地转录时，整个珠蛋白基因座的染色质结构发生改变（Weintraub and Groudine 1976；Stalder et al. 1980）。

致使基因座对 DNase 消化敏感的精确改变仍然未知，但是敏感性可能产生于基因座的解压缩，尽管事实上基因座在很大程度上仍然装配成核小体。是否所有的基因在表达细胞中比在非表达细胞中对核酸酶消化更敏感，或者对核酸酶消化的显著差异敏感性是否主要发生在多基因协调调控的基因座，如 β-珠蛋白基因、免疫球蛋白和 T 细胞受体基因座，这些也仍然不清楚。在我们的一个实验室里，针对一个分离基因的常规核酸酶敏感性研究，在基因活跃或是永久沉默的细胞类型间没有显示出有显著差异（M. Bradley and ST. Smale，未发表）。

在 1979 年，对 DNase I 敏感性的进一步研究导致了 DNase I 高敏感性的发现，后者与鉴定转录控制区更为相关。当利用果蝇热激基因进行 DNase 敏感性实验时，Carl Wu、Sarah Elgin 及其同事们以浓度逐渐增加的 DNase I 来消化细胞核。他们注意到，随着与活性基因相关的限制性片段在 Southern 印迹放射自显影上信号强度的下降，出现了比原限制性片段更小的“子带”（Wu et al. 1979a，b）。通过用不同的放射性标记 DNA 片段探测 Southern 印迹，并通过在 Southern 分析前用不同的限制酶切割 DNase 消化的基因组 DNA，将现在被称为 DNase I 高敏感性位点的“子带”定位于热激基因侧翼的特定区域（Wu 1980）（见专题 5.1）。

当在 20 世纪 70 年代晚期进行上述实验时，还未利用功能分析鉴定出远程控制区。但是，从这些数据推测：高敏感位点可能对应于在染色质结构中含有特别显著的改变的区域，该区域可能与转录控制区一致（Wu et al. 1979a，b；Stalder et al. 1980；McGinnis et al. 1983）。重要的是要注意，DNase 高敏感性和假定调控区之间的相互关系完全不同于在整个转录活跃的 β-珠蛋白基因座上观测到的 DNase 敏感性。

很快有许多研究确证了 DNase I 高敏感区对应于重要远程控制区。一个早期研究鉴定了在免疫球蛋白 κ 基因座内含子中的功能性增强子，该功能性增强子与以前所描述的 DNase I 高敏感位点一致（Picard and Schaffner 1984）。在其他类型的远程控制区中也已发现 DNase I 高敏感区，包括 LCR（Talbot et al. 1989）、沉默子（Sawada et al. 1994；Siu et al. 1994）、绝缘子（Udvardy et al. 1985）和 MAR（LevyWilson and Fortier 1989）。此外，正如早期的研究所证实的（Wu et al. 1979a，b；Wu 1980），启动子也能表现出对 DNase I 切割的高敏感性。

致使控制区对 DNase I 高敏感性的精确结构改变的机制还尚未确定，但由序列特异性 DNA 结合蛋白结合或者 ATP-依赖性核小体重塑复合物募集引起的核小体重塑或驱逐可能起到一定的作用。高分辨率的作图已揭示某些高敏感位点紧接着位于转录因子结合位点两侧，并且因此对应于由因子结合直接引起的结构改变。但是，控制区中的高敏感核苷酸不一定代表因子结合位点，因为当与核小体结合周边区域的敏感性相比较时，任何无蛋白质的核苷酸都有可能是高敏感性的。

虽然 DNase 高敏感性提供了 DNA 区域具有重要调控功能的证据，但要注意，已经鉴定出了似乎并不作为转录控制区起作用的高敏感性位点。这些位点的 DNase I 高敏感性可能是为了与功能无关的原因而存在。或者，可能需要改进的功能分析方法以揭示它们的功能。

用 DNase I 高敏感性鉴定控制区的优缺点

对目的基因内部和侧翼的 DNA 进行初步的系统发生比较后，强烈推荐将 DNase I 高敏感性作为鉴定远程控制区的出发点。通过使用各种不同的放射性标记探针（在基因序列信息基础上通过 PCR 制备）探测和再探测基因组 Southern 印迹，可迅速扫描跨越基因座的 20～100kb 或更长的区域。如果在基因座中发现了高敏感位点，通过数次另外的基因组 Southern 印迹，其中 DNase 消化的 DNA 在凝胶电泳前用不同限制酶彻底切割，就可以鉴定高敏感位点的精确位置。之后，通过 PCR 可以从基因组中扩增出含有高敏感性区域的 DNA 片段，再将扩增片段克隆进适当的载体中，并在功能分析中研究该 DNA 片段，以确定功能相关性。

作为 Southern 印迹方法的一种替代，已经发展了基于 PCR 的方法，以用于目的基因座中 DNase I 高敏感性位点的检测（McArthur et al. 2001；Follows et al. 2007）。另外，正如本章引言中所描述的，已经发展了用于 DNase I 高敏感性位点的全基因组鉴定的策略，并将这些策略用于鉴定越来越多的细胞类型中成千上万的高敏感性位点（Crawford et al. 2004，2006a，b；Dorschner et al. 2004；Sabo et al. 2004，2006；Follows et al. 2006；Boyle et al. 2008）。

对于常规 DNase I 高敏感性分析，用特定的放射性标记探针可检测 Southern 印迹上长达 20 kb 的限制性片段，以发现高敏感性位点的存在。但是，已经描述了一种非常有趣的策略，该策略利用电场反转凝胶电泳，以单个探针检验近 100 kb 的基因区域（Mautner et al. 1996；Pipkin and Lichtenheld 2006）。在常规 DNase I 高敏感性-Southern 印迹程序中，难度最大的步骤是基因座限制性作图和从基因组 DNA 克隆中分离恰当的杂交探针。但是，当许多生物的基因组序列信息面世，并且当制备探针的寡核苷酸（通过直接标记寡核苷酸或通过以寡核苷酸为 PCR 引物制备长的放射性探针）以最低成本容易获得时，这些挑战则被消除。有了这些进展，唯一剩下的技术挑战就是基因组 Southern 印迹上条带的检测，其中基因组 Southern 印迹是通过基因组 DNA 的 DNase I 不完全切割产生的，这些亚摩尔条带会很弱，并且难以检测。

DNase I 高敏感性方法的一个重要局限性是它不能鉴定基因的某些或全部相关控制区，因为并非所有的控制区都产生可鉴定的高敏感性位点。一个基因中所有重要控制区都已经被鉴定的最好证据是，装配一个合成基因座的能力，当导入转基因小鼠的种系时，该合成基因座能重现野生型水平的转录和精确调控。合成基因座可能含有远程控制区、启动子和该基因的转录单位。对于一些基因（如 β-珠蛋白基因）（Grosveld et al. 1987）来说似乎已实现了这一目标，但该结果并不具有决定性。为了组装合成构体，有必要删除介于远程控制区和启动子/转录单位之间的大段 DNA，或在其中插入大段无关 DNA。因为 DNA 的删除或插入可能改变基因座的活性，所以自信地确定一个基因的所有控制区已经全部被发现将会非常困难。

数据解释

当使用 Southern 印迹法时，DNase I 高敏感位点可以检测为明显的带或几乎不可检测的弱带。而且，有些带可能是清晰的，而有的带则可能是弥散的。在用含高敏感位点的 DNA 片段进行功能分析之前，不可能确定它们的功能相关性。

为了进一步表征高敏感性位点，可以分析表达或不表达该基因的其他细胞系和细胞类型。如果高敏感位点只在表达该基因的细胞类型中存在，该高敏感位点具有功能相关性的可能性就可能较大。和沉默子相联系的高敏感位点可能与该基因的表达呈负相关。但是遍在的高敏感位点对于调控也可能是关键的。例如，虽然由 LCR 调控的个别基因仅在部分发育阶段中表达，但是珠蛋白 LCR 中的高敏感性位点在来自红细胞发育所有阶段的细胞中都是可检测的（Crossley and Orkin 1993）。不管高敏感位点存在的细胞是何种类型，都将需要利用功能性方法确定该区域与调控性表达的相关性。

基质附着区的鉴定

核基质、MAR 和 SAR 的基本原理

MAR 或骨架结合区（SAR）的概念来源于高等真核细胞核内 DNA 构成大的染色质环的证据（Paulson and Laemmli 1977；Laemmli et al. 1992）。研究发现，每个环的基部与核骨架或基质的蛋白质组分相连（见第 1 章）。研究果蝇组蛋白和热激基因座时，Mirkovitch 等（1984）揭示了 DNA 与转录基因座附近特定的富含 AT 的序列相连，而并非随机地附着到基质上。之后，Gasser 和 Laemmli（1986）揭示某些 MAR（或 SAR）与转录控制区（如增强子）紧密结合。已鉴定和表征了来源于无脊椎动物和脊椎动物的很多 MAR，结果表明，许多但非全部 MAR 都紧靠转录调控序列（Laemmli et al. 1992）。

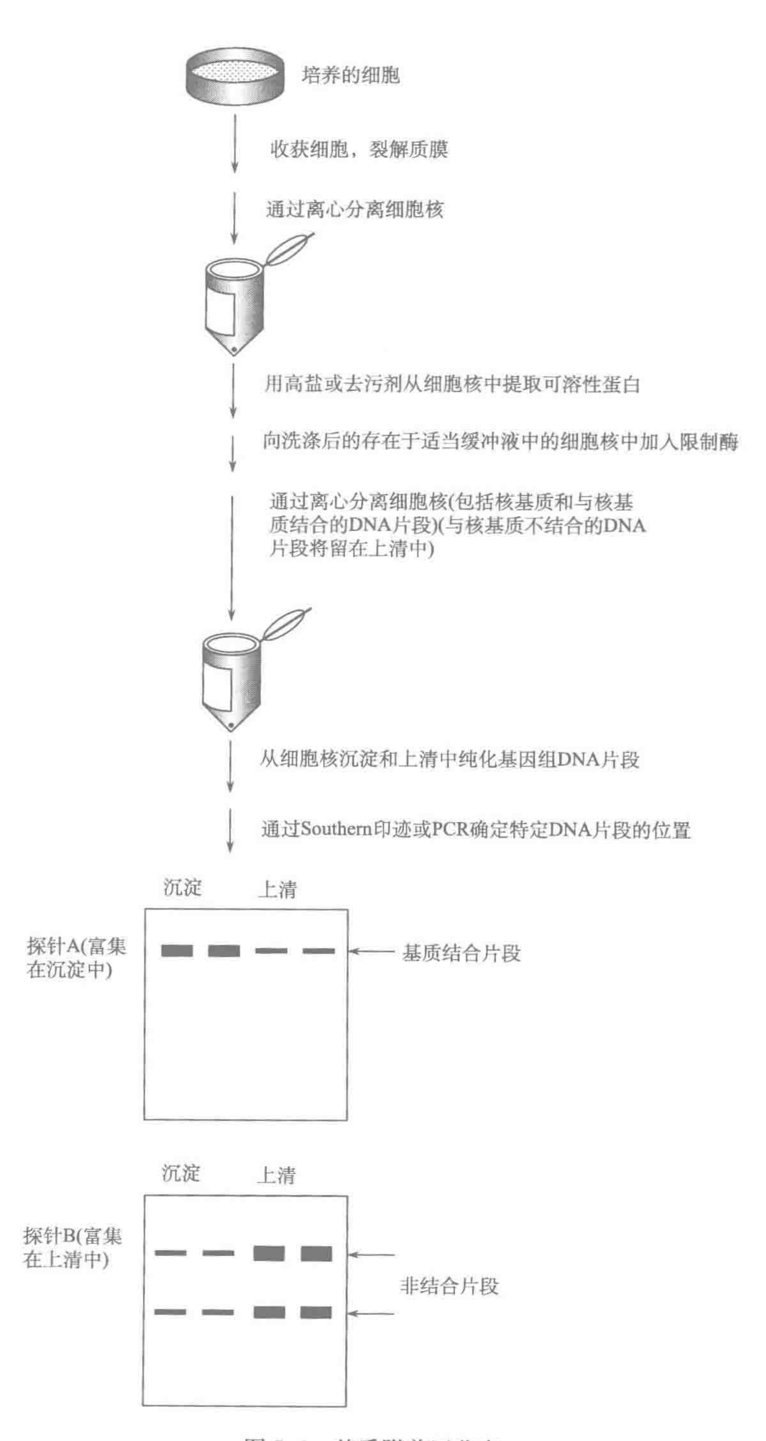
不幸的是，尚未明确确定基质结合的功能意义及 MAR 对基因调控发挥作用的机制（见第 1 章），这种不确定性的一个原因是不同的 MAR 似乎具有多种不同的功能。许多 MAR 都和转录增强子或 LCR 紧密相关（如见 Gasser and Laernmli 1986；Jarman and Higgs 1988），然而，已鉴定了在已使用的分析方法中不具有功能活性的 MAR（如见 Jarman and Higgs 1988）。

用 MAR 鉴定远程控制区的优缺点

作为鉴定远程控制区的一种方法，只有当本章中讨论的所有其他策略失败后，才推荐 MAR 的定位（见专题 5.2），该方法的主要局限性是 MAR 和远程控制区之间高度可变和不确定的关系。实际上可能只有一小部分的 MAR 与功能性重要控制区紧密相关。另外，大多数或所有与一个基因相联系的远程控制区都可能和 MAR 无关。尽管存在这种不确定性，基质结合的物理特征已成功地用于鉴定数个基因的远程控制区，并且如果其他策略失败，可以考虑使用这种方法。

专题 5.2　鉴定 MAR 的方法

用于鉴定 MAR 的方法（图 5.4）开始于利用高盐漂洗或去污剂分离细胞核并抽提



培养的细胞
收获细胞，裂解质膜
通过离心分离细胞核
用高盐或去污剂从细胞核中提取可溶性蛋白
向洗涤后的存在于适当缓冲液中的细胞核中加入限制酶
通过离心分离细胞核(包括核基质和与核基质结合的DNA片段)(与核基质不结合的DNA片段将留在上清中)
从细胞核沉淀和上清中纯化基因组DNA片段
通过Southern印迹或PCR确定特定DNA片段的位置
沉淀
上清
探针A(富集在沉淀中)
基质结合片段
沉淀
上清
探针B(富集在上清中)
非结合片段

图 5.4 基质附着区鉴定。

大部分可溶性蛋白质（包括组蛋白）(Cook and Brazell 1980；Mirkovitch et al. 1984)。然后利用限制性内切核酸酶切割基因组 DNA，并且通过离心和对不溶性沉淀物的洗涤，去除未与不溶性基质结合的 DNA 片段。利用一个或多个跨越目的基因座的放射性标记探针，通过 Southern 印迹分析与不溶性基质保持结合的 DNA 片段。对于确定 DNA 片段是否优先保留在不溶性基质中，PCR 是另一种切实可行的选择。与可溶性部分相比，DNA 片段在基质中的优先保留提示该片段含有真正的 MAR。在限制酶切割后，通过用 *Exo*III 消化基质部分，可进行 MAR 的精确定位。这种外切核酸酶可以从限制性片段的自由末端以 3′→5′ 的方向消化 DNA，但是却无法越过与基质蛋白质物理性结合的核苷酸。*Exo*III 处理后剩余片段的大小（通过 Southern 印迹分析确定）和与其杂交的放射性标记探针，揭示 MAR 的精确序列。利用这些标准程序鉴定的大多数 MAR 都富含 AT (Laemmli et al. 1992)。因此，如果鉴定为假定 MAR 的序列在 DNA 测序后发现富含 AT，则该结果将和以前的研究一致。

该程序中的唯一关键变量可能是用于制备核基质组分的方法，最初的方法是用高盐 (2 mol/L NaCl) 洗涤 (Cook and Brazell 1980)，该法可有效地从细胞核中除去了组蛋白和其他可溶性蛋白。但是，高盐洗涤也好像会改变蛋白质-DNA 的相互作用，至少有时会。为了避免这种情况，Mirkovitch 等 (1984) 发展了用去污剂二碘水杨酸锂代替 2 mol/L NaCl洗涤的基质制备方法。通过洗涤可很容易地将这种特别的去污剂从产生的基质制剂中去除，因而防止它干扰随后的限制酶切割。

鉴定远程控制区的功能性方法

功能性方法的基本优缺点

强烈推荐系统发生比较、DNase I 高敏感性实验，以及 ChIP 扫描方法作为鉴定远程控制区的出发点。但是，当假定控制区被定位后，需要用适当的功能性分析方法确定其是否确实功能相关。功能性方法也可以替代或者与上面提到的分析方法组合使用，以鉴定和定位远程控制区。这些功能性方法有同时解决两个挑战的优点：鉴定潜在的调控区，并为该区域发展功能性分析方法。此外，功能性策略有可能鉴定出“非功能性”方法遗漏的关键控制区。这些策略的一个主要缺点是：如果没有通过简单的系统发生分析或者 DNase I 高敏感性对潜在的控制区定位，那么这些策略可能耗费更多的时间和工作量。在大多数情况下，这一缺点超过优点。

始于基因组 DNA 大片段的功能性方法

利用功能性分析鉴定远程控制区的两种主要方法本质上是相互对立的。第一种方法从测试包含或者围绕目的基因的基因组 DNA 大片段开始（图 5.5），它比第二种方法更值得推荐。通过稳定转染可以将基因组片段目的性地导入转基因小鼠或培养的细胞，然后监测目的基因的转录，以确定 DNA 片段是否含有正确调控所需的所有控制区。如果观测到正确调控，使用相同的分析方法测试各种亚片段，以确定重要控制区的位置。

就此方法而言，恰当的出发点是获得含有目的基因的黏粒、P1 或细菌人工染色体 (BAC) 克隆。由于已注释的 BAC 文库可轻易获取，所以对哺乳动物基因来说，目前

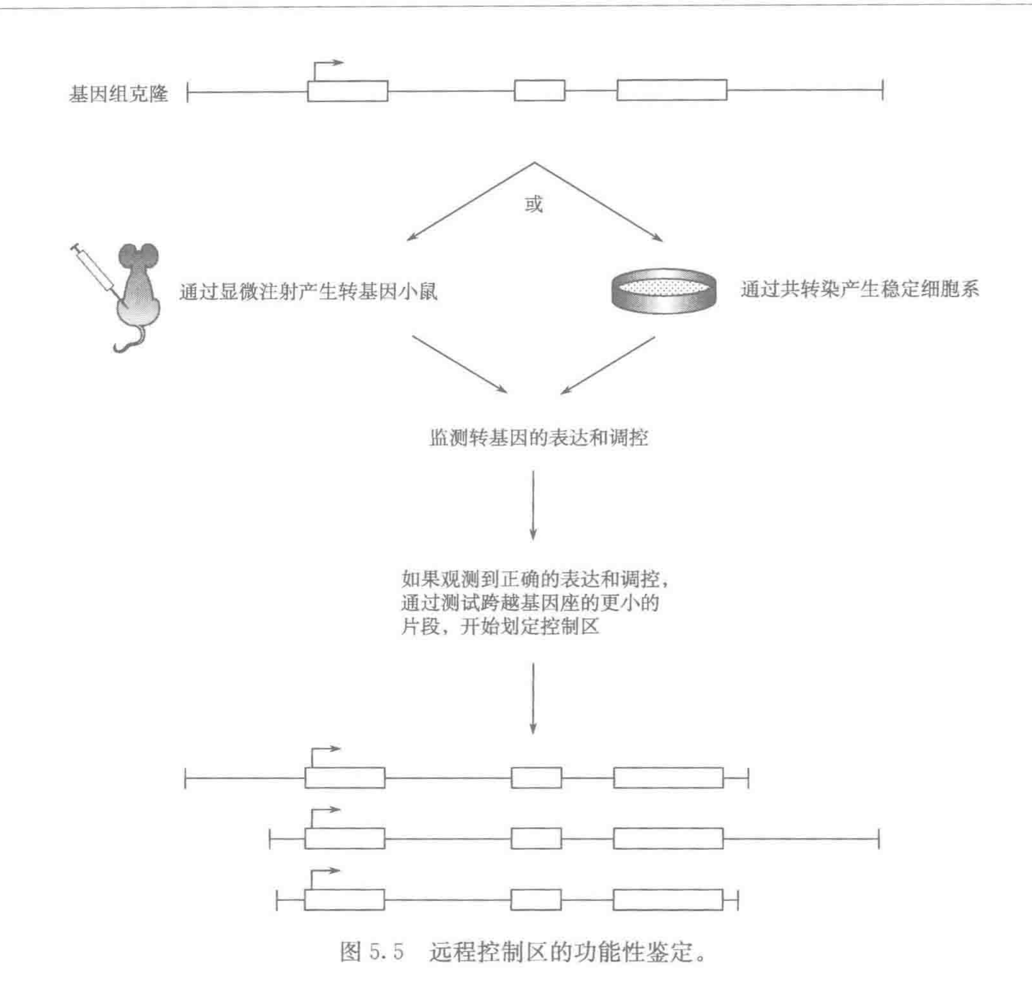


图 5.5 远程控制区的功能性鉴定。

强烈推荐 BAC 克隆（如 BACPAC Resource Center at Children's Hospital Oakland Research Institute，http://bacpac.chori.org/ 或 Invitrogen BAC/PAC Genomic DNA Resources，http://clones.invitrogen.com/cloneinfo.php? clone=bacpac)。在 UCSC 基因组浏览器（http://genome.ucsc.edu）上可鉴定包含有目的 DNA 区域的单个 BAC，并且可以最低成本购买，无需文库筛选（Sparwasser et al. 2004）。强烈推荐 BAC 也是因为已经发展了在大肠杆菌中通过同源重组制作 BAC 克隆的高效策略，这些策略消除了以前使用标准分子克隆技术修饰 DNA 大片段的困难。特别是，Heintz 和他的同事们开发出一种经优化用于大肠杆菌（*E. coli*）中 BAC 修饰的穿梭载体（Yang et al. 1997；Misulovin et al. 2001；Gong et al. 2002；Testa et al. 2004)。Copeland、Jenkins 及其同事发展了一种称为 BAC 重组工程的较新的策略，当在 BAC 中引入突变时，这种策略提供了一些有利条件（http://recombineering.ncifcrf.gov/）（Warming et al. 2005）。

常用的方法是在大肠杆菌中使用同源重组将报告基因［如荧光素酶或绿色荧光蛋白（GFP）基因］导入目的基因的转录单位中。报告基因使得容易地监控目的基因的转录成为可能，也使区分来自 BAC 转基因的转录本和来自内源性基因的转录本成为可能

[早期研究是将人基因导入转基因小鼠中用来区分内源性基因和转基因（如 Blom van Assendelft et al. 1989；Strouboulis et al. 1992），但因为现在将报告基因或者小的序列标签导入 DNA 大片段非常简单，不再推荐跨物种实验，而且调控区在不同物种中不可能发挥最佳作用]。然后将修饰后的 BAC 注射进卵母细胞的原核里，以产生转基因小鼠品系（Yang et al. 1997；Testa et al. 2004）。可监测报告基因的活性，以确定 BAC 转基因是否被正确调控，从而提供所有重要的控制区都位于 BAC 中的证据。也可以用定量反转录 PCR 来比较来自 BAC 和目的内源性基因的表达，其中 PCR 使用与报告基因中的序列退火的引物。如果观察到 BAC 转基因有效地表达和正确调控，就可以用同样的转基因分析来产生和分析亚片段。

使用这种方法进行的一个最有趣和最完整的分析是，鉴定和表征调控 *Rag1* 和 *Rag2* 基因表达的控制区（Yannoutsos et al. 2004）。该研究分析了转基因小鼠中大约 30 个不同的 DNA 片段，以鉴定负责 *Rag1* /*Rag2* 正确调控的区域。该研究首先证明来自 BAC 转基因的转录被正确调控，其中 BAC 转基因含有 *Rag1*/*Rag2* 基因座，并在该基因座的 *Rag2* 转录单位中插入 GFP 报告基因。连续的含有较小的亚片段的转基因分析鉴定了有助于控制 *Rag1*/*Rag2* 表达的沉默子和远程反义沉默子。对于每个被测验的构体，有必要分析多个创建小鼠以确保转基因整合位点不严重影响表达。这是有可能的，因为转基因创建者是在出生后不久被分析的，未繁殖或尝试维持创建系。

另外一个有趣的例子与小鼠 T-辅助者 2（Th2）细胞因子基因簇的分析有关，该基因簇包含编码 IL-4、IL-5 和 IL-13 的基因。因为这个基因座对于调节免疫应答具有十分重要的作用，为了搜索远程控制区，已经用系统发生分析和 DNase I 高敏感性分析对其进行了广泛的表征。这些方法鉴定了数个关键调控区，并且当通过同源重组从基因组中删除内源性区域时，可牢固确立这些区域的功能（Nardone et al. 2004）。但是，当 Flavell 及其同事使用 BAC 转基因方法鉴定正确调控所需的最小 DNA 区域时，发现以往研究中鉴定的区域是不足的（Lee et al. 2003）。相反，正确调控需要有一个位于组成型表达基因 RAD50 中的区域，而 RAD50 基因就嵌入在 Th2 细胞因子基因簇中。在系统发生和 DNase I 高敏感性研究中，这一重要区域会被遗漏，其中的原因部分是因为这些研究盯着基因间区域，而不是与 RAD50 基因紧密相连的区域。BAC 转基因结果强烈暗示这一新的调控区作为一个 LCR 起作用，因为对于转基因创建者稳定驱动从 *Il4* 基因座的高效转录，它是必需的；同时，对于转录水平与转基因拷贝数大约成正比，它也是必需的。通过同源重组从内源性基因中随后删除该 LCR 的一部分，确证了它对 Th2 细胞因子正确调控的重要性（Lee et al. 2005）。

从非常大的 DNA 片段开始的优势在于，在初始试验中观察到正确调控基因表达的可能性相对较高。如果是这样，研究者已成功地发展了切实可行的并且与生物学相关的功能性分析方法，且将处于开始定位重要的远程控制区的状态。在基因组序列信息获得之前，为非常大的 DNA 片段确定限制性图谱非常困难，以至于这种困难抵消了这个优势。但是，对于基因组已经测序的许多生物来说，这一挑战已经被消除了。此前，DNA 大片段很难用标准分子克隆技术修饰。通过同源重组在大肠杆菌（*E. coli*）中操纵 BAC 的策略的出现已经消除了这一主要挑战。无需生成和筛选文库，就可以以最低

的成本购买含目的基因的 BAC，代表了另一个主要进展。

一般通过产生转基因小鼠，而不是通过导入已建立的细胞系进行 BAC 的研究；由于它们的尺寸大，BAC 难以通过转染导入细胞内。相比之下，通过卵母细胞中 DNA 的显微注射产生转基因小鼠，则消除了对转染的需要。此外，筛选已经在基因组中稳定整合了 BAC 的细胞，一般需要基于药物抗性的选择，并且一直难以将这样的抗药性基因引入 BAC 中 [通过发展一种以新霉素抗性基因改造 BAC 的简单策略，这一障碍最近已被消除（Wang et al. 2006）]。由于显微注射 DNA 分子整合到卵母细胞基因组的频率高，所以对于转基因小鼠的产生来说，抗药性基因并不是必需的。最后，首选转基因小鼠，因为得到的结果比用细胞系得到的结果更有可能是生物相关的。然而，产生和分析转基因小鼠相关的时间及成本对很多实验室来说是主要障碍。

值得一提的是关于使用 BAC 分析基因调控的另外两种思想。第一，虽然 BAC 很难通过转染进入一般的转化细胞系中，但是它们却能非常容易地转染进小鼠胚胎干（ES）细胞（Testa et al. 2004）。因为 ES 细胞在培养中可以分化成各种不同的细胞类型，所以为了避免产生具有被修饰的 BAC 的转基因小鼠，一种可能的策略是将修饰的 BAC 通过转染导入到 ES 细胞中（以抗药基因改造后），然后将 ES 细胞分化成目的细胞类型。

第二，尽管强烈推荐 BAC 转基因方法作为鉴定和表征目的基因远程控制区的初始步骤，并且据认为 BAC 转基因方法是实现这一目标的一种高度的生理性策略，但是 BAC 转基因有时容易受染色体整合位点变化的影响。因此，BAC 转基因可能并不总是表征内源性基因调节机制的理想策略；对目的内源性基因直接操控将是更好的策略。一种设想鉴定重要远程控制区的策略是，从内源性基因座中系统地删除显示高度进化保守性的远程区域。在过去，此类系统性方法是难以考虑的，因为序列保守性的详细信息一般是难以获得的，而且还存在从内源性基因座中删除保守区所需要的时间、成本和工作量等问题。但是，现在详细的序列信息可以轻易获得，并且技术的进展使得同源重组也已经比较容易进行。另外，对于远程控制区的系统分析，已经发展了便于将多种不同突变导入目的基因座的策略（Wu et al. 1994；van der Weyden et al. 2002）。结合用于生成打靶构体的 PCR 策略的进展，此类用于鉴定和表征远程控制区的系统分析正迅速变得可行。

始于指导报告基因表达的较小片段的功能性方法

上述方法的代替方法是一种扫描策略（图 5.6）。可将许多跨越基因座的 5～10 kb 的较小片段导入含有来自目的基因的启动子或异源启动子 [如猿猴病毒 40（SV40）或单纯疱疹病毒胸苷激酶（HSV-TK）启动子] 的报告载体中。然后，可用瞬时转染或稳定转染分析方法对所产生的一组质粒进行测验。如果其中一个质粒含有增强报告基因活性的片段，则可以进一步测验较小的亚片段，以便更精确地定位相关控制区。随着基因组序列信息的获得，另外一种可能性是专门测验表现进化保守性的远程区，而不是扫描目的基因所有侧翼区。

用扫描法鉴定和定位的一个增强子是 T 细胞受体 α 增强子（Ho et al. 1989；Winoto and Baltimore 1989）。用 DNase I 高敏感性鉴定增强子的最初的努力失败了。作为一种替代方法，可以将跨越基因座的 DNA 片段插入到氯霉素乙酰转移酶报告质粒上，并在转化的 T 细胞系内进行瞬时转染分析测验（Winoto and Baltimore 1989）。一

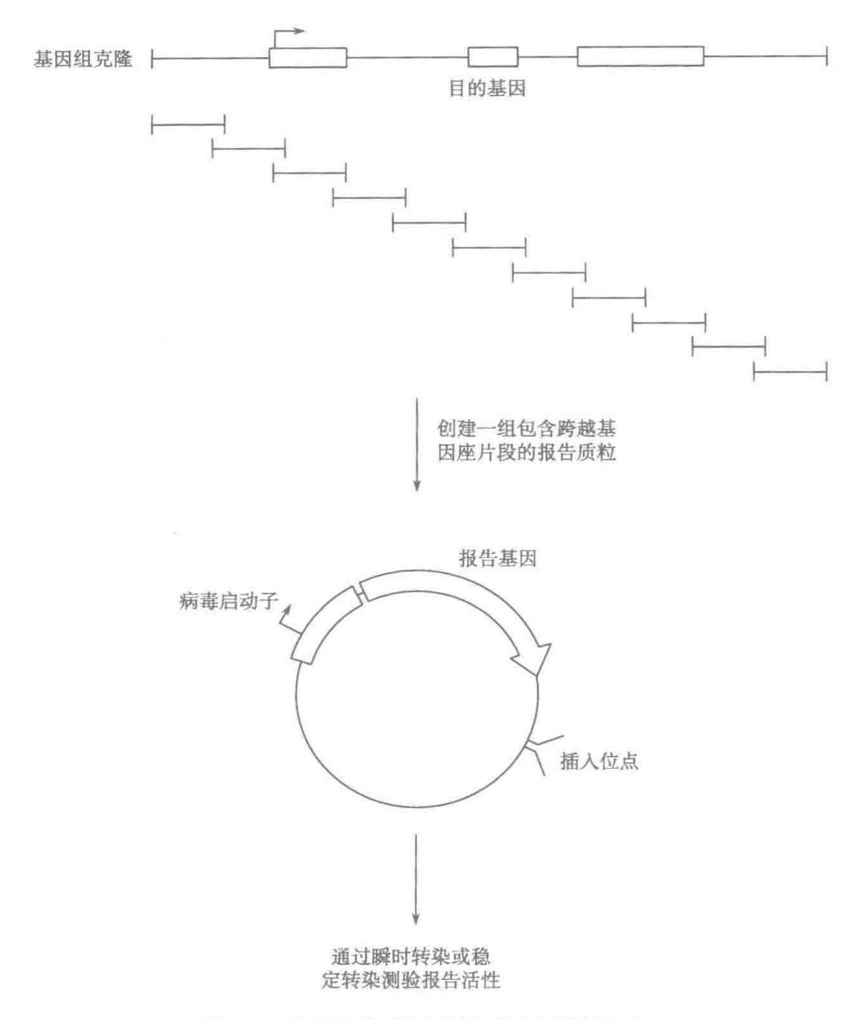


图 5.6　通过扫描的远程控制区功能性鉴定。

个质粒，包含基因下游约 5 kb 的一个 DNA 区域，产生了大幅度增强的报告基因活性。对来源于这一区域的亚片段的分析精确定位了这一增强子。

该方法存在以下几个优点。第一，灵敏的报告基因可以很容易地用于确定 DNA 片段何时增强或减弱基因转录。第二，鉴于可以容易地通过 PCR 将片段克隆进简单的报告质粒中，该分析可非常快速地进行。第三，质粒一般含有小片段的基因组 DNA，如果检测到活性，可快速鉴定功能性区域。第四，通过简单的瞬时转染分析，可以检测许多控制区（如增强子、某些 LCR 和某些沉默子）的活性，这可能更符合实验室的资源及能力。

这种策略的主要缺点是，在瞬时转染甚至稳定转染分析中 DNA 小片段的分析代表一种高度人工化的方法。因为 DNA 片段是在染色质环境下检测可能的功能，这种环境显著偏离内源性基因相关的自然环境，所以遗失 DNA 片段重要功能的可能性很大。同样，DNA 片段在转染分析中支持增强子或沉默子活性的可能性很大，而这些活性与内源性基

因的正确调控无关。与之相关的一个缺点是，这种分析可能不能为目的基因揭示任何关键控制区。相比之下，虽然BAC转基因方法需要更多的时间和工作量，但是这种方法最终获得重要新信息的可能性更大。因为这个原因，BAC转基因的应用更值得推荐。

表征远程控制区的功能性分析方法

瞬时转染分析

如果通过系统发生分析或者DNase I高敏感性、ChIP或MAR实验鉴定了一个假定控制区，就需要一种分析方法以评价其相关性。在哺乳动物细胞中，研究人员必须首先在瞬时转染分析、稳定转染分析、转基因分析或者内源性干扰分析（通过同源重组在基因组中消除假定控制区）中选择一种分析方法。这里也提到了与瞬时转染和稳定转染分析相关的其他考虑。

第4章描述了分析启动子的策略，其中介绍了为分析远程控制区而发展一种瞬时分析方法需要考虑的大部分问题。首先需要适当的细胞资源和适当的转染条件。另外，必须选择报告基因和分析方法。对于启动子分析，重要的是确证启动子中的控制元件从正确的核苷酸处指导转录。对于远程控制区的分析，因为远程控制区不直接确定起始位点的定位，所以起始位点的定位与远程控制区的分析相关性不大。因此，起始位点的变化对远程控制区分析结果的解释不太重要。

必须将报告质粒本身设计得不同于用于检测启动子活性的报告质粒。通常，在大多数市售报告载体中，将目的DNA片段插入位于报告基因3′端的限制性位点中。在这些载体中，在报告基因的上游一般包含病毒启动子，这些启动子介导目的DNA片段赋予的活性（见第4章）。因为在少数例子中远程控制区与它自己的启动子相配合时可以最有效地发挥作用（Li and Noll 1994；Merli et al. 1996；Ohtsuki et al. 1998），因此用来自目的基因的启动子代替病毒启动子可能是值得的。但是，因为绝大多数增强子、LCR和沉默子对异源启动子起作用，所以这种替代没有必要，除非用别的方法都难以检测活性时，才作为最后的对策。

目的片段通常插入到报告基因的3′端。能在瞬时分析中起作用的大多数远程控制元件都应该易于从这一位置上发挥功能。更重要的是，如果目的片段插入到启动子的最近上游，即使DNA片段不是一个真实的增强子或沉默子，细胞核蛋白的偶然结合位点也可能会与启动子中的邻近控制元件发生协同作用并影响报告基因的活性。

正如启动子分析中所描述的，瞬时转染实验以比较简单的方式进行（见第4章）。为了归一化，通常在转染实验中包含一个适当的内部对照。如果分析涉及不同细胞系之间的比较，可使用其他对照质粒来归一化转染效率，其中对照质粒含有与相同报告基因融合的病毒启动子和增强子。归一化会有助于确立控制区是否真正赋予细胞类型特异的调控。

稳定转染分析

实质上稳定转染分析可按第4章中所描述的进行。简而言之，用两种质粒共转染进入细胞，其中一种质粒含有目的控制区调控的报告基因，另一种质粒含有病毒启动子/增强子控制的抗药性基因。这两种质粒通常以二者的多拷贝共整合到基因组的一个或多个位点上。最后通过加入药物选择抗药性基因稳定整合的细胞。抗药性细胞中通常也包

含稳定整合的报告质粒。如果一个 DNA 片段不管是插入到报告基因的上游还是下游，都能一贯地增强稳定整合的启动子-报道基因活性，那么可以将该 DNA 片段界定为增强子。

当分析诸如增强子和 LCR 的远程控制区时，必须考虑的一点是驱动抗药性基因表达的病毒增强子能掩盖目的控制区对报告基因转录的刺激，这是因为病毒增强子可能会作用于定位在报告基因上游的启动子。可以将绝缘子元件置于报告基因盒的任何一端，以减少或消除调控耐药基因的控制区影响报告基因转录的可能性（Chung et al. 1993; Hsu et al. 2003）（见第 4 章）。添加绝缘子元件也可以降低染色体整合位点对报告基因转录的影响。在 *Rag1* 和 *Rag2* 基因的新增强子分析中，报道质粒上添加鸡 β-珠蛋白绝缘子（Chung et al. 1993）极大地增强了依赖这一新增强子的转录激活的幅度，并降低了整合位点的变化性（Hsu et al. 2003）。

降低抗药性基因盒对报告基因转录影响的另一策略是，使用相对弱的启动子来驱动抗药性基因的转录，这种方法促进了许多珠蛋白 LCR 研究的成功（如见 Blom van Assendeift et al. 1989）。在这种方法中，将抗药性基因置于一个无增强子的 HSV-TK 启动子控制之下。作为另一种替代方法，可以考虑在第 4 章中曾简要提及的一种方法。可让目的控制区直接调控抗药基因（或致癌基因），以取代共转染病毒启动子/增强子调控的抗药基因和目的控制区调控的报告基因（图 5.7）（如见 Fromm and Berg 1982; Zhong and Krangel 1997）。将异源启动子置于抗药性基因的上游，并将用于分析增强子、沉默子或 LCR 活性的 DNA 片段插入抗药性基因下游。可将产生的质粒稳定转染进适当的细胞，并进行药物筛选。抗药性克隆的数目与调控抗药性基因转录的控制区强度应大致成正比。因此，如果插入的 DNA 片段作为增强子或 LCR 起作用，应该获得大量的抗药性克隆，其数目超过启动子单独调控抗药性基因时获得的抗药性克隆数目。对于贴壁型细胞，通过在标准的组织培养皿中接种和选择细胞，以及对生长在确定区域中的克隆进行计数，可确定抗药性克隆的数目。对于非贴壁细胞，通过将转染细胞接种到 96 孔板中，然后对选择后含有克隆的孔进行计数，以确定抗药性克隆的数目。

LCR 活性的表明

上述介绍的策略对研究增强子特别有用，这些增强子在所有已讨论的分析方法（瞬时转染、稳定转染及转基因分析）中都起作用。但是，其他类型的远程控制区，如 LCR、沉默子和绝缘子通常只在这些分析方法中的部分分析方法中起作用。另外，可能需要进行特定类型的分析，以表明远程控制区符合 LCR、沉默子或绝缘子的定义。例如，要确凿地表明一个调控区起 LCR 的作用，需要转基因小鼠分析。LCR 最简单的定义是能将染色体整合位点无关的高水平表达赋予其连接基因的 DNA 片段（Fraser and Grosveld 1998）（见第 1 章）。LCR 还刺激 mRNA 合成，其合成水平与整合的转基因拷贝数大致成正比。因为 LCR 的这些定义特征依赖于染色体整合，因此不能使用瞬时转染分析，尽管有些 LCR 在瞬时分析中起增强子作用（Fraser and Grosveld 1998）。

为了利用稳定转染分析鉴定一个控制区为 LCR，必须分离数个稳定克隆，这些克隆含有受目的区调控的报告基因。如果相当数量的稳定转染克隆都具有一个整合的报告基因和控制区，但不能产生可检测的报告基因活性，则控制区不可能是 LCR。相比之

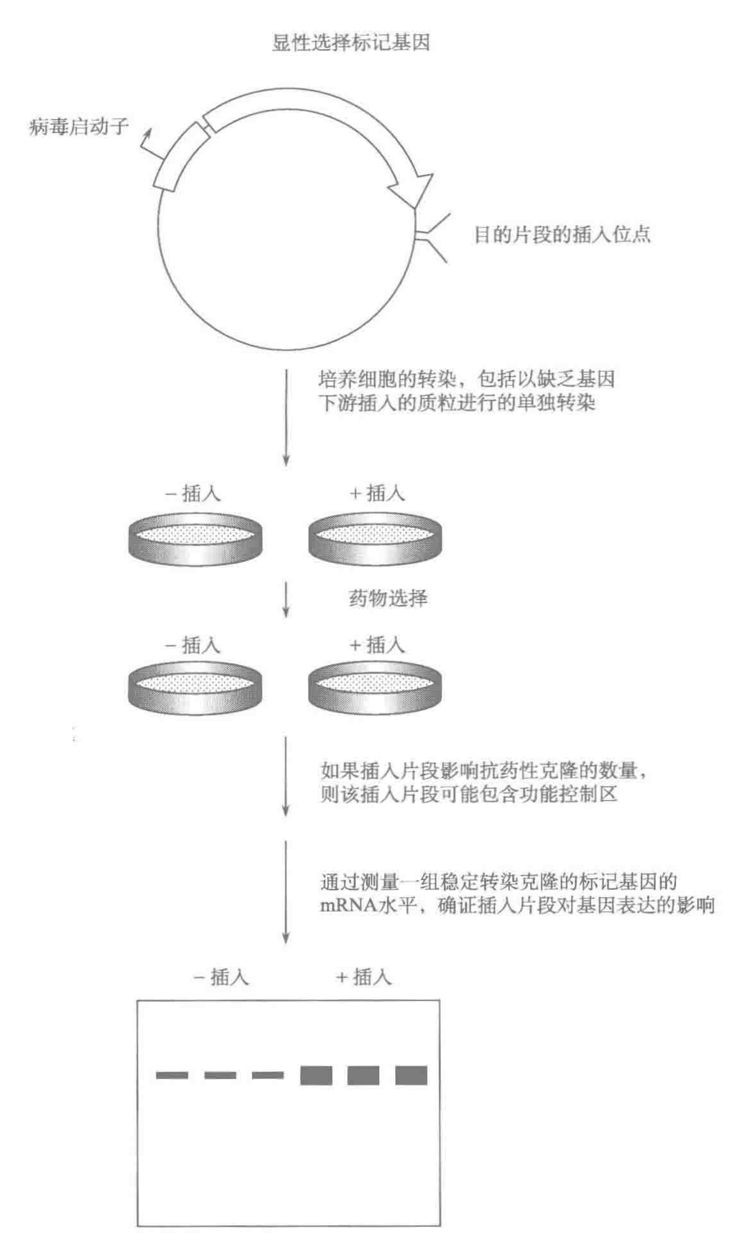


图 5.7 通过监测显性选择标记基因的调控进行远程控制区的功能性鉴定。

下，如果获得了一贯高水平的报告基因活性，则该结果将支持这样一个假设，即该区域是一个 LCR。但该结果仍然是非决定性的，因为除非外源 DNA 整合进一个至少部分是“可及的”染色体位置，否则将不能获得抗药性克隆。因为抗药性基因和报告基因几乎总是共整合进相同的位点（见第 4 章），因此在抗药性细胞中，报告基因通常整合进可及的位置。鉴于这个原因，即使控制区不是真正的 LCR，也可能会在抗药性克隆中观

察到一贯高水平的报告基因活性。于是，尽管已经用稳定的分析方法研究了 LCR 活性 (Blom van Assendelft et al. 1989)，但是这种分析不能提供一个控制区起 LCR 作用的令人信服的证据。

如上所述，转基因分析是唯一能令人信服地证明一个控制区具有 LCR 关键特征 (即高水平的整合位点无关表达) 的分析方法。将含有目的控制区的报告质粒显微注射进小鼠卵母细胞，将产生完全随机的染色体整合。因为不涉及药物筛选，随后的结果不偏向于进入可及位置的整合体。可以分析数个含有种系整合报告基因的创建株，以确定在相关组织中获得的表达水平。如果该控制区是一个 LCR，事实上在所有已经整合了完整质粒的创建系中，将会检测到高水平的报道活性，且表达水平和转基因拷贝数大致成正比。相比之下，如果远程控制区是缺乏 LCR 典型特征的增强子，则只能在相对小部分的创建系中检测到高水平的报告基因活性，剩下的创建系一般表现出低水平的报道活性或者无活性。

沉默子活性的表明

术语“沉默子”实际上用于界定与增强子和 LCR 相对立的元件。更具体地说，沉默子是在所有分析或只在一部分确定的分析中抑制基因表达的远程控制区。关于沉默子在其中起作用的分析方法的类型还存在歧义，即使对于来自鼠 CD4 基因座的经典沉默子也是如此。Siu 及其同事已经报道，需要进行转基因小鼠分析，以检测一贯的和令人信服的 CD4 沉默子活性 (Siu et al. 1994; Duncan et al. 1996)。特别是，在转基因小鼠中，沉默子抑制了处于 CD4 启动子和增强子控制下的报告基因的转录。相比之下，Sawada 等 (1994) 报道，当沉默子以多重串联拷贝存在于典型的启动子和增强子上游时，它在瞬时转染分析中起作用。最有可能的是，瞬时转染产生能反映沉默子某些特征的相对适中的沉默子活性。但是，要真正理解沉默子功能的机制及其对染色质结构的可能影响，转基因分析 (或许稳定转染分析) 可能是必需的。因为尚不清楚研究沉默子功能需要哪种类型的分析，所以从转基因小鼠分析开始是最有意义的。该区域被鉴定后，可以尝试稳定转染和瞬时转染分析，以确定它们对后续研究是否有用。

绝缘子活性的表明

绝缘子的定义特征是它隔离或者保护启动子免受上游增强子影响的能力。仅有少数基因的绝缘子已经被描述，并且基于核酸酶高敏感性、基质附着、功能及其在核酸敏感区和抗性区之间边界上的存在，已经鉴定了这些绝缘子 (Kellum and Schedi 1991; Chung et al. 1993; Kalos and Fournier 1995; Gdula et al. 1996)。有些绝缘子作为附加质粒起作用 (Dunaway et al. 1997)，但大多数似乎需要染色体整合。因为这一领域的研究尚未充分发展，所以最好参考最初的文献以确定如何鉴定并分析一个新的绝缘子。要检验一个控制区的绝缘子功能，通常将其插入到报告质粒的增强子和启动子之间 (Kellum and Schedl 1992)。重要的是要表明只有当该区域定位于增强子和启动子之间，而非定位于二者上游时，该区域才能抑制报告基因的转录。

封闭是绝缘子的基本特性，虽然依此特征表征了几乎所有已界定的绝缘子，但是有两种绝缘子特征呈现不同。第一，大多数绝缘子不具有转录激活能力，仅能封闭远程控制区的功能。相比之下，至少有一个控制区 (和载脂蛋白 B 基因相连) 已被描述为绝

缘子，该绝缘子也显示了与 LCR 的特征相重叠于的特征（Kalos and Fournier 1995）。这种绝缘子最初被鉴定为一种边界元件，这种边界元件位于核酸酶敏感性载脂蛋白 B 基因座与核酸酶抗性旁侧区之间。这一特征暗示它能使旁侧区绝缘于基因座内发生的染色质改变，反之亦然。实际上，在转基因小鼠中，边界区使报告基因绝缘于整合位点的影响。但是，这种绝缘子也增强报告基因活性，和 LCR 类似。这一区域与真正的 LCR 的明显区别是 LCR 一般在两个方向上增强核酸酶敏感性，而载脂蛋白区则充当核酸酶敏感性与抗性之间的边界。

绝缘子还在它们阻断其活性控制区的类型上存在不同。例如，和果蝇 87A7 基因相联系的绝缘子阻断增强子的活性，但不能阻断 LCR 的活性（Kellum and Schedl 1992）。相反，在鸡 β-珠蛋白基因座 5′ 边界发现的绝缘子阻断增强子和 LCR 两者的活性（Chung et al. 1993）。当设计方法以测定假定绝缘子区域的功能时，必须考虑到这些可变的特征。

总之，不管一个控制区是通过系统发生比较、DNase I 高敏感性、ChIP 分析、基质附着实验，还是通过功能性方法鉴定，对于功能性研究来说，恰当出发点是通过在各种分析中测试一个控制区，界定与该区域相关的活性及特征的范围。例如：

- 它在瞬时转染分析中起作用吗？
- 它是只在自己的启动子存在时起作用，还是在异源启动子存在时同样具有活性？
- 它在稳定转染分析中起作用吗？
- 它是以整合位点依赖的方式起作用吗？
- 它使启动子绝缘于远程增强子的影响吗？
- 它与核基质结合吗？
- 它对应于明显的 DNase I 高敏感区吗？它与无核小体区或核小体重建区相一致吗？

通过分析活性和特征的范围，有可能对正在研究的区域的类型作出有把握的判断。一个重要的结果将是关于恰当的功能性分析方法的决定，以用于后续分析。后续分析的目标是：①鉴定该区域内的重要控制元件；②阐明该区域调控转录的机制。

参考文献

Barski A., Cuddapah S., Cui K., Rob T. Y., Schones D. E., Wang Z., Wei G., Chepelev I., and Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* 129: 823-837.

Bejerano G., Siepel A. C., Kent W. J., and Haussler D. 2005. Computational screening of conserved genomic DNA in search of functional noncoding elements. *Nat. Methods* 7: 535-545.

Blom van Assendelft G., Hanscombe O., Grosveld F., and Greaves D. R. 1989. The β-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. *Cell* 56: 969-977.

Boyle A. P., Davis S., Shulha H. P., Meltzer P., Margulies E. H., Weng Z., Furey T. S., and Crawford G. E. 2008. High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132: 311-322.

Chung J. H., Whiteley M., and Felsenfeld G. 1993. A 5′ element of the chicken β-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. *Cell* 74: 505-514.

Cook P. R. and Brazell I. A. 1980. Mapping sequences in loops of nuclear DNA by their progressive detachment from

the nuclear cage. *Nucleic Acids Res*. 8: 2895-2906.

Crawford G. E., Davis S., Scacheri P. C., Renaud G., Halawi M. J., Erdos M. R., Green R., Meltzer P. S., Wolfsberg T. G., and Collins F. S. 2006a. DNase-chip: A high-resolution method to identify DNase I hypersensitive sites using tiled arrays. *Nat. Methods* 3: 503-509.

Crawford G. E., Holt I. E., Mullikin I. C., Tai D., Blakesley R., Bouffard G., Young A., Masiello C., Green E. D., Wolfsberg T. G., and Collins F. S. 2004. Identifying gene regulatory elements by genome-wide recovery of DNase hypersensitive sites. *Proc. Natl. Acad. Sci.* 101: 992-997.

Crawford G. E., Holt I. E., Whittle J., Webb B. D., Tai D., Davis S., Margulies E. H., Chen Y., Bernat J. A., Ginsburg D., et al. 2006b. Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). *Genome Res*. 16: 123-13 1.

Crossley M. and Orkin S. H. 1993. Regulation of the β globin locus. *Curr. Opin. Genet. Dev.* 3: 232-237.

de Last W., Klous P., Kooren J., Noordermeer D., Palstra R. J., Simonis M., Splinter E., and Grosveld F. 2008. Three-dimensional organization of gene expression in erythroid cells. *Curr. Top. Dev. Biol.* 82: 117-139.

Dekker J., Rippe K., Dekker M., and Kleckner N. 2002. Capturing chromosome conformation. *Science* 295: 1306-1311.

Dorschner M. O., Hawrylycz M., Humbert R., Wallace J. C., Shafer A., Kawamoto J., Mack J., Hall R., Goldy J., et al. 2004. High- throughput localization of functional elements by quantitative chromatin profiling. *Nat. Methods* 1: 219-225.

Dunaway M., Hwang J. Y., Xiong M., and Yuen H. L. 1997. The activity of the scs and scs' insulator elements is not dependent on chromosomal context. *Mol. Cell. Biol.* 17: 182-189.

Duncan D. D., Adlam M., and Siu G. 1996. Asymmetric redundancy in CD4 silencer function. *Immunity* 4: 301-311.

Elgin S. R. and Weintraub H. 1975. Chromosomal proteins and chromatin structure. *Annu. Rev. Biochem.* 44: 725-774.

Enver T., Brewer A. C., and Patient R. K. 1985. Simian virus 40-mediated cis induction of the Xenopus β-globin DNase I hypersensitive site. *Nature* 318: 680-683.

Follows G. A., Janes M. E., Valuer L., Green A. R., and Gottgens B. 2007. Real-time PCR mapping of DNase I-hypersensitive sites using a novel ligation-mediated amplification technique. *Nucleic Acids Res*. 35: e56.

Follows G. A., Dhami P., Gottgens B., Bruce A. W., Campbell P. J., Dillon S. C., Smith A. M., Koch C., Donaldson I. J., Scott M. A., et al. 2006. Identifying gene regulatory elements by genomic microarray mapping of DNase I hypersensitive sites. *Genome Res*. 16: 1310-1319.

Fraser P. and Grosveld F. 1998. Locus control regions, chromatin activation and transcription. *Curr. Opin. Cell Biol*. 10: 36 1-365.

Fromm M. and Berg P. 1982. Deletion mapping of DNA regions required for SV40 early region promoter function in vivo. *J. Mol. Appl. Genet.* 1: 457-481.

Gasser S. M. and Laemmli U. K. 1986. Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of D. melanogaster. *Cell* 46: 521-530.

Gdula D. A., Gerasimova T. I., and Corces V. G. 1996. Genetic and molecular analysis of the gypsy chromatin insulator of Drosophila. *Proc. Natl. Acad. Sci.* 93: 9378-9383.

Gong S., Yang X. W., Li C., and Heintz N. 2002. Highly efficient modification of bacterial artificial chromosomes (BACs) using novel shuttle vectors containing the R6Kgamma origin of replication. *Genome Res*. 12: 1992-1998.

Grass J. A., Boyer M. E., Pal S., Wu J., Weiss M. J., and Bresnick E. H. 2003. GATA- 1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proc. Natl. Acad. Sci.* 100: 8811-8816.

Grosveld F., van Assendelft G. B., Greaves D. R., and Kollias G. 1987. Position-independent, high-level expres-

sion of the human β-globin gene in transgenic mice. *Cell* 51: 975-985.

Heintzman N. D., Stuart R. K., Hon G., Fu Y., Ching C. W., Hawkins R. D., Barrera L. O., Van Calcar S., Qu C., Ching K. A., et al. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* 39: 311-318.

Hinrichs A. S., Karolchik D., Baertsch R., Barber G. P., Bejerano G., Clawson H., Diekhans M., Furey T. S., Harte R. A., Hsu F., et al. 2006. The UCSC Genome Browser Database: Update 2006. *Nucleic Acids Res.* 34: D590-D598.

Ho I. C., Yang L. H., Morle G., and Leiden J. M. 1989. A T-cell-specific transcriptional enhancer element 3′ of C alpha in the human Tcell receptor alpha locus. *Proc. Natl. Acad. Sci.* 86: 6714-6718.

Hsu L. Y., Lauring J., Liang H. E., Greenbaum S., Cado D., Zhuang Y., and Schlissel M. S. 2003. A conserved transcriptional enhancer regulates RAG gene expression in developing B cells. *Immunity* 19: 105-117.

Jarman A. P. and Higgs D. R. 1988. Nuclear scaffold attachment sites in the human globin gene complexes. *EMBO J.* 7: 333 7-3344.

Kalos M. and Fournier R. E. 1995. Position-independent transgene expression mediated by boundary elements from the apolipoprotein B chromatin domain. *Mol. Cell. Biol.* 15: 198-207.

Karolchik D., Bejerano G., Hinrichs A. S., Kuhn R. M., Miller W., Rosenbloom K. R., Zweig A. S., Haussler D., and Kent W. J. 2007. Comparative genomic analysis using the UCSC genome browser. *Methods Mol. Biol.* 395: 17-34.

Karolchik D., Kuhn R. M., Baertsch R., Barber G. P., Clawson H., Diekhans M., Giardine B., Harte R. A., Hinrichs A. S., Hsu F., et al. 2008. The UCSC Genome Browser Database: 2008 update. *Nucleic Acids Res.* 36: D773-D779.

Kellum R. and Schedl P. 1991. A position-effect assay for boundaries of higher order chromosomal domains. *Cell* 64: 941-950.

Kellum R. and Schedl P. 1992. A group of scs elements function as domain boundaries in an enhancer-blocking assay. *Mol. Cell. Biol.* 12: 2424-2431.

Kent W. J., Sugnet C. W., Furey T. S., Roskin K. M., Pringle T. H., Zahler A. M., and Haussler D. 2002. The Human Genome Browser at UCSC. *Genome Res.* 12: 996-1006.

Laemmli U. K., Kas E., Poijak L., and Adachi Y. 1992. Scaffold-associated regions: cis-acting determinants of chromatin structural loops and functional domains. *Curr. Op. Genet. Dev.* 2: 275-285.

Lee G. R., Spilianakis C. G., and Flavell R. A. 2005. Hypersensitive site 7 of the TH2 locus control region is essential for expressing TH2 cytokine genes and for long-range intrachromosomal interactions. *Nat. Immunol.* 6: 42-48.

Lee G. R., Fields P. E., Griffin T. J., and Flavell R. A. 2003. Regulation of the Th2 cytokine locus by a locus control region. *Immunity* 19: 145-153.

Levy-Wilson B. and Fortier C. 1989. The limits of the DNase I-sensitive domain of the human apolipoprotein B gene coincide with the locations of chromosomal anchorage loops and define the 5′ and 3′ boundaries of the gene. *J. Biol. Chem.* 264: 21196-21204.

Li X. and Noll M. 1994. Compatibility between enhancers and promoters determines the transcriptional specificity of gooseberry and gooseberry neuro in the Drosophila embryo. *EMBO J.* 13: 400-406.

Mautner J., Bornkamm G. W., and Polack A. 1996. Long-range chromatin analysis of the human MYC locus by pulsed-field gel electrophoresis. *Genes Chromosomes Cancer* 16: 247-253.

McArthur M., Gerum S., and Stamatoyannopoulos G. 2001. Quantification of DNase I-sensitivity by real-time PCR: Quantitative analysis of DNaseI-hypersensitivity of the mouse globin LCR. *J. Mol. Biol.* 313: 27-34.

McGinnis W., Shermoen A. W., Heemskerk J., and Beckendorf S. K. 1983. DNA sequence changes in an upstream DNase I-hypersensitive region are correlated with reduced gene expression. *Proc. Natl. Acad. Sci.* 80:

1063-1067.

Merli C., Bergstrom D. E., Cygan J. A., and Blackman R. K. 1996. Promoter specificity mediates the independent regulation of neighboring genes. *Genes Dev.* 10: 1260-1270.

Mirkovitch J., Mirault M. E., and Laemmli U. K. 1984. Organization of the highest-order chromatin loop: Specific DNA attachment sites on nuclear scaffold. *Cell* 39: 223-232.

Misulovin Z., Yang X. W., Yu W., Heintz N., and Meffre E. 2001. A rapid method for targeted modifications and screening of recombinant bacterial artificial chromosomes. *J. Immunol. Methods* 257: 99-105.

Nardone J., Lee D. U., Ansel K. M, and Rao A. 2004. Bioinformatics for the "bench biologist": How to find regulatory regions in genomic DNA. *Nat. Immunol.* 5: 768-774.

Ohtsuki S., Levine M., and Cai H. N. 1998. Different core promoters *possess* distinct regulatory activities in the Drosophila embryo. *Genes Dev.* 12: 547-556.

Paulson J. R. and Laemmli U. K. 1977. The structure of histonedepleted metaphase chromosomes. *Cell* 12: 817-828.

Picard D. and Schaffner W. 1984. A lymphocyte-specific enhancer in the mouse immunoglobulin K gene. *Nature* 307: 80-82.

Pipkin M. E. and Lichtenheld M. G. 2006. A reliable method to display authentic DNase I hypersensitive sites at long-ranges in single-copy genes from large genomes. *Nucleic Acids Res.* 34: e34.

Roh T. Y. and Zhao K. 2007. High-resolution, genome-wide mapping of chromatin modifications by GMAT. *Methods Mol. Biol.* 387: 95-108.

Rob T. Y., Cuddapah S., and Zhao K. 2005. Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes Dev.* 19: 542-552.

Rob T. Y., Cuddapah S., Cui K., and Zhao K. 2006. The genomic landscape of histone modifications in human T cells. *Proc. Natl. Acad. Sci.* 103: 15782-15787.

Rob T. Y., Wei G., Farrell G. M., and Zhao K. 2007. Genome-wide prediction of conserved and nonconserved enhancers by histone acetylation patterns. *Genome Res.* 17: 74-81.

Sabo P. J., Humbert R., Hawrylycz M., Wallace J. C., Dorschner M. D., McArthur M., and Stamatoyannopoulos J. A. 2004. Genome-wide identification of DNase I hypersensitive Sites using active chromatin sequence libraries. *Proc. Natl. Acad. Sci.* 101: 4537-4542.

Sabo P. J., Kuehn M. S., Thurman R., Johnson B. E., Johnson E. M., Gao H., Yu M., Rosenzweig E., Goldy J., Haydock A., et al. Genomescale mapping of DNase I sensitivity in vivo using tiling DNA microarrays. *Nat. Methods* 3: 511-518.

Sawada S., Scarborough J. D., Killeen N., and Littman D. R. 1994. A lineage-specific transcriptional silencer regulates CD4 gene expression during T lymphocyte development. *Cell* 77: 917-929.

Schones D. E. and Zhao K. 2008. Genome-wide approaches to studying chromatin modifications. *Nat. Rev. Genet.* 9: 179-191.

Siepel A., Bejerano G., Pedersen J. S., Hinrichs A. S., Hou M., Rosenbloom K., Glawson H., Spieth J., Hillier L. W., Richards S., et al. 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* 15: 1034-1050.

Simonis M., Kooren J., and de Laat W. 2007. An evaluation of 3C-based methods to capture DNA interactions. *Nat. Methods* 4: 895-901.

Siu G., Wurster A. L., Duncan D. D., Soliman T. M., and Hedrick S. M. 1994. A transcriptional silencer controls the developmental expression of the CD4 gene. *EMBO J.* 13: 3570-3579.

Smale S. T. 2003. A Th cytokine LCR. Adding a new piece to the regulatory puzzle. *Immunity* 19: 1-2.

Sparwasser T., Gong S., Li J. Y., and Eberl G. 2004. General method for the modification of different BAG types and the rapid generation of BAG transgenic mice. *Genesis* 38: 39-50.

Spilianakis G. G. and Flavell R. A. 2004. Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat. Immunol.* 5: 1017-1027.

Spilianakis C. G., Lalioti M. D., Town T., Lee G. R., and Flavell R. A. 2005. Interchromosomal associations between alternatively expressed loci. *Nature* 435: 637-645.

Stalder J., Larsen A., Engel J. D., Dolan M., Groudine M., and Weintraub H. 1980. Tissue-specific DNA cleavages in the globin chromatin domain introduced by DNase I. *Cell* 20: 451-460.

Strouboulis J., Dillon N., and Grosveld F. 1992. Developmental regulation of a complete 70-kb human β-globin locus in transgenic mice. *Genes Dev.* 6: 1857-1864.

Szutorisz H., Canzonetta C., Georgiou A., Chow G. M., Tora L., and Dillon N. 2005. Formation of an active tissue-specific chromatin domain initiated by epigenetic marking at the embryonic stem cell stage. *Mol. Cell. Biol.* 25: 1804-1820.

Talbot D., Collis P., Antoniou M., Vidal M., Grosveld F., and Greaves D. R. 1989. A dominant control region from the human β-globin locus conferring integration site independent gene expression. *Nature* 338: 352-355.

Testa G., Vintersten K., Zhang Y., Benes V., Muyrers J. P. P., and Stewart A. F. 2004. BAC engineering for the generation of ES cell-targeting constructs and mouse transgenes. *Methods Mol. Biol.* 256: 123-139.

Tolhuis B., Palstra R. J., Splinter E., Grosveld F., and de Laat W. 2002. Looping and interaction between hypersensitive sites in the active β globin locus. *Mol. Cell* 10: 1453-1465.

Udvardy A., Maine E., and Schedl P. 1985. The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. *J. Mol. Biol.* 185: 341-358.

van der Weyden L., Adams D. J., and Bradley A. 2002. Tools for targeted manipulation of the mouse genome. *Physiol. Genomics* 11: 133-164.

Wang Z., Engler P., Longacre A., and Storb U. 2006. An efficient method for high-fidelity BAC/PAC retrofitting with a selectable marker for mammalian cell transfection. *Genome Res.* 11: 137-142.

Warming S., Costantino N., Court D. L., Jenkins N. A., and Copeland N. G. 2005. Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res.* 33: e36.

Weinmann A. S., Plevy S. E., and Smale S. T. 1999. Rapid and selective remodeling of a positioned nucleosome during the induction of IL-12 p40 transcription. *Immunity* 11: 665-675.

Weintraub H. and Groudine M. 1976. Chromosomal subunits in active genes have an altered conformation. *Science* 193: 848-856.

Winoto A. and Baltimore D. 1989. A novel, inducible and T cell-specific enhancer located at the 3′ end of the T cell receptor alpha locus. *EMBO J.* 8: 729-733.

Wu C. 1980. The 5′ ends of Drosophila heat shock genes in chromatin are hypersensitive to DNase I. *Nature* 286: 854-860.

Wu C., Wong Y. -C., and Elgin S. C. R. 1979a. The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activity. *Cell* 16: 807-814.

Wu H., Liu X., and Jaenisch R. 1994. Double replacement: Strategy for efficient introduction of subtle mutations into the murine Colla-1 gene by homologous recombination in embryonic stem cells. *Proc. Natl. Acad. Sci.* 91: 2819-2823.

Wu C., Bingham P. M., Livak K. J., Holmgren R., and Elgin S. C. 1979b. The chromatin structure of specific genes: I. Evidence for higher order domains of defined DNA sequence. *Cell* 16: 797-806.

Yang X. W., Model P., and Heintz N. 1997. Homologous recombination based modification in Escherichia coli and germline transmission in transgenic mice of a bacterial artificial chromosome. *Nat. Biotechnol.* 15: 859-865.

Yannoutsos N., Barreto V., Misulovin Z., Gazumyan A., Yu W., Rajewsky N., Peixoto B. R., Eisenreich T., and Nussenzweig M. C. 2004. A cis element in the recombination activating gene locus regulates gene ex-

pression by counteracting a distant silencer. *Nat. Immunol.* 5：443-450.

Zhong X. -P. and Krangel M. S. 1997. An enhancer blocking element between α and δ gene segments within the T cell receptor α/δ locus. *Proc. Natl. Acad. Sci.* 94：5219-5224.

Zhou L.，Nazarian A. A.，Xu J.，Tantin D.，Corcoran L. M.，and Smale S. T. An inducible enhancer required for *Il12b* promoter activity in an insulated chromatin environment. *Mol. Cell. Biol.* 27：2698-2712.

网络资源

http://bacpac.chori.org/：在儿童医院奥克兰研究所（the Children's Hospital Oakland Research Institute）的BACPAC资源中心（The BACPAC Resource Center），具有目前可用的BAC和PAC基因组DNA文库，BAC、PAC及F黏粒（fosmid）克隆，cDNA集合，高密度克隆杂交膜，以及BAC和PAC克隆载体等方面的信息。

http://clones.invitrogen.com/cloneinfo.php? clone=bacpac：Invitrogen公司的BAC/PAC基因组DNA资源主页。

http://ensembl.org：对于鉴定和评价真核基因组中保守的非编码区，Ensembl基因组浏览器（The Ensembl Genome Browser）是有用的工具。

http://genome.ucsc.edu：对于鉴定和评价真核基因组中保守的非编码区，加利福尼亚大学圣克鲁斯基因组浏览器（The University of California，Santa Cruz Genome Browser）是有用的工具。

http://genome.ucsc.edu/ cgi-binlhgTrackUi? hgsid=105645537&c chr20&g=dulceDnaseCd4Sites：人CD4⁺T细胞中全基因组DNase I高敏感性位点的汇编。

http://recombineering.ncifcrf.gov/：该网站提供了重组介导的遗传工程有关的详细的方案、参考文献，以及质粒图谱和序列。

6 鉴定控制区内的顺式作用 DNA 元件

要点

• 计算方法可用于鉴定进化上保守的 DNA 序列模体，以及对应于已知 DNA 结合蛋白识别位点的模体。当与染色质免疫沉淀及其他基础实验相结合时，这些计算方法能提供对调控目的基因的元件及因子的初步认识。

• 当追求对目的控制区更加完整的了解时，一般应进行综合性突变体分析，以此作为进行更高级的功能研究的关键步骤。

引言和概述

前面几章描述了鉴定转录控制区，包括启动子和各种远程控制区的方法及策略。为了着手阐明控制区调控转录的机制，接下来必须描述相关的顺式作用 DNA 序列元件并界定与那些元件相互作用的反式作用蛋白质因子。在本书的第一版中，我们强烈建议研究人员在尝试鉴定相关的转录因子前，鉴定目的控制区内重要 DNA 元件。我们进一步推荐了将综合性突变体分析作为鉴定目的控制区内重要 DNA 序列元件的必不可少的第一步。当时，还没有建立一种可靠的、监测活细胞内特异性蛋白质与目的控制区相结合的技术，所以有关蛋白质结合的研究一般在体外进行，还需要大量的工作来确定同一蛋白质是否与控制区在体内也结合，以及在生理条件下，这样的结合对转录控制区的功能是否重要。我们还推荐了综合性突变体分析，它可用于鉴定可能对基因的调控起作用的 DNA 模体的计算策略（基于进化上的保守性以及已知转录因子识别位点的存在），但在当时才刚刚起步。

然而，监测体内蛋白质-DNA 相互作用的染色质免疫沉淀（ChIP）技术已有巨大进展，而 DNA 序列的系统发生分析已有更大的进展。因此，现在我们建议研究人员一开始就对控制区进行系统发生分析，以鉴定出在进化中一直强烈保守的 DNA 序列模体（见第 5 章）。因为已经获得了如此大量的生物的完整或基本完整的基因组序列，所以模体的强烈保守性（如在所有脊椎动物的进化中）提供了比以前所能提供的更强大的功能相关性证据，以前只有两个或三个物种可以比较。在鉴定出进化中保守性好的模体后，就能对每个模体进行计算分析（见第 7 章和下文），以确定它们是否对应于已知蛋白质或蛋白质家族的一致（consensus）结合位点或近一致（near-consensus）结合位点。如

果有候选蛋白的抗体，就能进行 ChIP 实验，以确定该候选蛋白是否在体内与该控制区有联系。如果结果是肯定的，就能启动功能性研究，以确定该蛋白质是否对该控制区的活性起作用。功能性研究一般包括转录因子功能缺失实验或功能获得实验，或者破坏该蛋白质预测 DNA 结合位点的实验（见第 8 章）。

虽然这个策略能够洞悉大量的调控目的基因的转录因子，但是也有许多值得注意的警告事项，从而使得我们推荐它的时候有些犹豫。首先，最初的系统发生分析只会鉴定出那些进化中保守性好的 DNA 元件。系统发生分析会漏掉那些必然存在于某些控制区内的物种特异性 DNA 元件。遗憾的是，特定基因的调控策略和关键 DNA 元件在不同物种间变化的程度仍旧未知。如果将进化保守作为鉴定重要元件的唯一标准的话，就将会被遗漏的关键调控元件的百分比而论，研究人员的预测千差万别。此外，许多 DNA 结合蛋白能适应 DNA 识别序列的大幅度改变而不降低结合亲和力，这进一步增加了系统发生分析至少会漏掉一些关键元件的可能性。

一个有关的并且更大的顾虑是，用于预测哪个已知 DNA 结合蛋白与一个 DNA 模体发生相互作用的计算程序依然是原始的。如第 2 章所述，这些程序的质量主要取决于用于生成它们的实验数据的质量。一个显而易见的问题是蛋白质组中只有少数 DNA 结合蛋白的一致识别位点已通过实验确定。而且，对于任何真核 DNA 结合蛋白来说，尚没有其识别序列范围的全面信息，并要求蛋白质与该范围内所有识别序列的结合要有足够的亲和力以产生功能活性。与此问题相关的是，某些蛋白质通过低亲和力位点发挥功能，这些低亲和力位点与一致识别序列有相当大的不同。这些蛋白质一般与其他 DNA 结合蛋白协同性地结合。最后一个问题是，许多 DNA 结合蛋白都是多蛋白大家族的成员，其中多个家族成员识别相似或相同的 DNA 序列。因此，即使计算程序成功预测出哪个蛋白质家族能够与一个重要的进化保守的 DNA 模体结合，也可能难以预测该家族中哪个成员（或哪些成员）在其自然基因组环境中负责该模体的功能。由于这些局限性，以及预测哪些蛋白质与进化保守的 DNA 模体结合的其他局限性，用计算的方法很可能不会揭示出相关的 DNA 结合蛋白。在这些情况下，可能有必要利用第 7 章概述的方法之一及第 8 章描述的评价功能相关性的策略，实验鉴定能与保守元件相结合的蛋白质。

在我们推荐的策略中，下一步是用 ChIP 实验检查候选蛋白在体内结合目的控制区的能力。重要的是要牢记 ChIP 不是一种功能性分析（见第 8 章和第 9 章），因此，证明蛋白质在体内结合控制区仅为这种相互作用的功能相关性提供了有限的认识。ChIP 进一步受限于它对目的蛋白的高质量抗体和适于 ChIP 分析的蛋白质-DNA 相互作用的依赖。由于种种原因，包括瞬时结合或与 DNA 的低效交联或表位屏蔽（第 8 章和第 9 章），一些蛋白质难以用 ChIP 分析来进行研究。

尽管有这些严重警告，系统发生分析和 ChIP 实验现在仍然被推荐为功能性研究的起始步骤，因为它们能快速提供有用的知识。如果研究的目标是为目的控制区鉴定部分调控元件和 DNA 结合蛋白，并且如果对计算方法建议的候选蛋白有足够的兴趣，研究者可能希望将未来的研究重点集中在这些元件和蛋白质上，而不对控制区进行更全面的分析。此外，如果研究者致力于完全理解某个基因的调控，那么强烈推荐在合适的功能性分析（如利用启动子-报告质粒进行瞬时转染）中进行综合性突变体分析，以鉴定所

有对控制元件的功能起作用的元件。这种分析将洞悉在所选的分析方法中哪些进化保守的元件对控制区活性是关键的，以及在系统发生分析中关键 DNA 元件是否因为进化中不够保守而被遗漏了。不过，正如下文所描述的，综合性突变体分析也有它自己重大的局限性。

我们不建议研究者在没有通过系统发生分析或诱变首先精确定位可能在功能上重要的单个 DNA 模体位置的情况下，就简单地用生化方法鉴定与目的控制区在体外结合的蛋白质。众所周知，核提取物包含大量能与任何 100～400 bp（一般启动子或增强子的长度）DNA 片段结合的蛋白质。因此，必然能找到与目的 DNA 片段结合的蛋白质，但是利用此方法优先鉴定出相关 DNA 结合蛋白的可能性很低。

要牢记，不管研究者是否从系统发生分析开始导向 ChIP 实验，还是从综合性突变体分析开始导向 ChIP 实验，获得的结果都将对调控该基因的机制只提供初步的看法。可能需要利用第 7 章所描述的实验策略去鉴定能与已鉴定的 DNA 模体结合的蛋白质。而且，事实上对于所有研究，正如第 8 章所描述的，将有必要仔细评价每个蛋白质-DNA 相互作用的功能相关性。

实验策略

通过综合性突变体分析鉴定控制元件

如上所述，目前我们推荐研究者从系统发生分析、已知 DNA 结合蛋白识别模体的计算分析，以及可能的 ChIP 实验开始，获得有关负责目的控制区功能的控制元件和因子的初步知识。在第 5 章和第 7 章中分别简单地讨论系统发生分析和对识别模体的分析，在第 8 章和第 9 章中将讨论 ChIP，在本章中，我们将重点讨论综合性突变体分析的首选策略。本书包括该讨论的原因是，当新进入转录领域的研究人员尝试进行这种分析时，很多详细的技术问题的答案通常不明显。我们也注意到进行综合性突变体分析的频率不高，因为研究人员将其重点转向系统发生分析、ChIP 分析，以及转录因子功能缺失研究。尽管如此，由于这些方法各自都有很大的局限性（第 8 章），最终将需要详细的突变体分析来实现对调控任何真核基因相关机制的深入理解。

重要的是要强调，虽然综合性突变体分析具有上文所述的潜在的优点，但同时它也有三个局限性。第一，一次有意义的分析比系统发生分析或 ChIP 分析需要投入更多的精力。所幸的是，随着亚克隆技术和体外诱变技术的改进，综合性突变体分析所需的时间正在不断减少。第二，更大的局限性在于所鉴定到的 DNA 元件仅包含控制区在所选用的分析中发挥作用所需的那些元件。瞬时转染分析可以鉴定控制区发挥作用所需的大多数元件。然而，如第 4 章和第 5 章所讨论的，控制元件可能存在于仅在自然染色质环境中才起作用的启动子或增强子内。而且，瞬时转染后经常存在于细胞中的高质粒拷贝数可能导致一些重要的调控元件被忽略，或较不重要的调控元件占优势。稳定转染分析可以有更好的机会揭示控制区内所有重要的元件。然而，当用于综合性突变体分析时，稳定转染分析是极其耗时的，而且也可能无法揭示自然基因组环境中所需的所有 DNA 模体的功能。综合性突变体分析的第三个局限性与控制区内 DNA 元件可能存在的冗余性有

关。如果冗余性存在，就难以确定控制区的边界，进而难以鉴定重要的 DNA 元件。

难以权衡综合性突变体分析的优点和缺点。一方面，综合性分析可以揭示重要的调控元件，这些元件在对已知蛋白的识别位点进行的系统发生分析或计算分析中被遗漏了。另一方面，突变分析需要投入更多的精力，而且所获得的信息很可能仍不全面。因此，必须针对研究的总体目标，再结合可用的实验资源及时间要求，对该问题进行评估（见第 2 章和上文）。

综合性分析的策略

大多数基本的诱变研究都是将缺失和置换突变体并用。例如，我们描述一下对 *Il12b* 基因启动子的解析，鼠源基因 *Il12b* 编码细胞因子白细胞介素-12（IL-12）和白细胞介素-23（IL-23）的 p40 亚基（Plevy et al. 1997）。我们首先关注用于勾画经胞外刺激诱导后 *Il12b* 启动子活性所需的 DNA 元件的策略；这种入手方式能应用于任何控制区，不管其调控方式如何。本章稍后将讨论哪些元件直接负责诱导。

要解析 *Il12b* 启动子，首先要进行缺失分析，主要是为了界定功能性启动子的边界。然后，通过使用一系列每个 6～10 bp 的成簇置换突变对该区域进行扫描，鉴定在那些边界内的单个控制元件。不使用小的缺失的原因是它们有可能改变控制元件的排列，这种排列对于正确调控可能是重要的。之后是对每个控制元件进行 3 bp 的精细置换突变，以确定它们的边界，并确定每个控制元件是否代表了一个蛋白质的结合位点或 2～3 个蛋白质的混合位点。下面介绍设计策略中的每一步时要考虑的问题，以及每一步所获得的信息。虽然这个实例是针对启动子的解析，但相似的策略同样可望用于增强子和其他远程控制区。

Il12b 基因在细菌产物，如脂多糖（LPS）活化的巨噬细胞和树状突细胞中表达。该分析最初的目的是通过 LPS 解析基因诱导的机制。因为已表征的大多数可诱导细胞因子基因主要受启动子序列调控，所以可能认为 *Il12b* 启动子将负责诱导作用，至少部分地负责诱导作用。在我们的分析之前，Trinchieri 及其同事用核连缀分析（第 2 章）表明诱导作用主要在转录起始水平上被调控（Ma et al. 1996），当时，鼠源基因的转录起始位点已经定位在富 AT 序列下游大约 25 bp 处，其中富 AT 序列可能作为 TATA 框起作用。

在标准报告载体（Plevy et al. 1997）中，一个 405 bp 的启动子片段（相对于起始位点从－350 延伸至＋55）与萤光素酶和氯霉素乙酰转移酶（CAT）报告基因相融合。一般不推荐 CAT 报告基因，因为 CAT 报告分析比其他报告分析如萤光素酶（见第 4 章）更耗时，而灵敏性却不及其他报告分析。但是，在该研究中发现 CAT 报告基因却更可取，因为此时可获得的萤火虫萤光素酶 cDNA 含有导致萤光素酶活性在 LPS 刺激巨噬细胞中的非特异性诱导（Plevy et al. 1997）。对于 *Il12b* 启动子插入来说，选择－350作为上游边界，是因为大多数启动子在距起始位点几百个碱基对内都有关键元件。＋55 处的下游边界靠近翻译起始密码子，选择＋55 作为下游边界，是因为有些基因在非翻译前导序列中含有重要的启动子元件（见第 4 章）。将这些质粒转染到 RAW 264.7 巨噬细胞系中并经 LPS 活化后，发现 405 bp 片段足以产生强的、可诱导的启动子活性

(Plevy et al . 1997)。因此，需要使用诱变策略来鉴定该片段内重要的控制元件。

缺失分析

突变体分析的第一步是生成一系列 5′ 缺失突变体并对其进行分析（Plevy et al. 1997)。使用跨越缺失预期终点的正向引物和跨越核苷酸＋55 的反向引物，通过聚合酶链反应（PCR）制备缺失。两个引物都含有与启动子片段终点相邻的限制性位点序列，从而使 PCR 产物经内切核酸酶切割后，可以直接插入报告载体。因为突变是通过 PCR 产生的，所以有必要对最终的质粒插入片段进行测序，以确保没有产生不必要的点突变。

在制备置换突变体前要先制备缺失突变体并对其加以分析，因为缺失分析可界定支持完整活性的最小序列。通过确定最小序列，随后必须制备的置换突变体的数量就能减少到最低。例如，如果整个 405 bp 片段是诱导活性所必不可少的，就需要 41 个 10 bp 置换突变扫描该区域的重要元件。相反，如果该片段 5′ 端的 200 bp 去除后对活性没有重要影响，那就只需要 21 个 10 bp 置换突变对该功能相关区域进行扫描。事实上，结果揭示出－200 至＋55 的片段保留有 100％的启动子活性，其中－150～＋ 55 的片段保留了 40％的启动子活性（图 6.1)。此外，当序列延伸到－800 时，启动子活性没有显著增强。

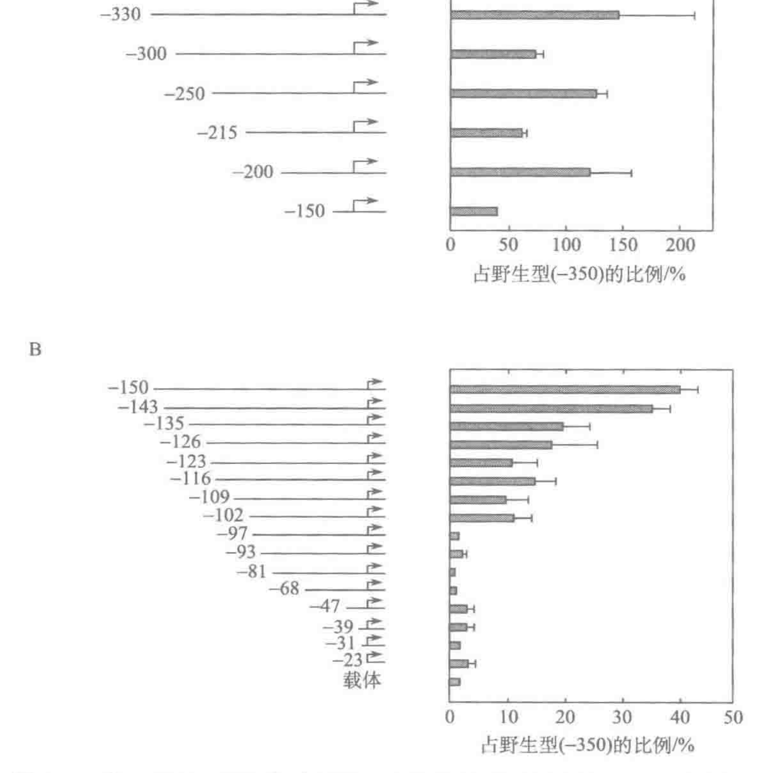


图 6.1 鼠源基因 *Il12b* 启动子的基本缺失分析（经允许，改编自 Plevy et al. 1997，© American Society for Microbiology)。

关于这些初步缺失分析结果的解释，还有几点值得注意。首先，图 6.1 的数据表明 −350 和 −150 之间的序列缺失对启动子活性影响不大。例如，−250～−215 的缺失体使启动子活性降低了两倍，−215～−200 的缺失体使启动子活性增强了两倍，而 −200～−150 的缺失体又再次使启动子活性降低了两倍。导致这些效应的缺失序列中可能含有重要的正、负控制元件，或者两倍差异与启动子活性无关。每个缺失体都构建成 *Il12b* 基因座序列与载体序列的融合体。当载体序列与启动子内不同的核苷酸相融合时，载体序列对启动子活性多少会有一些或正或负的影响。因此，不能简单地推断每个两倍差异都对应一个相关的启动子元件。例如，为了确定 −250 和 −215 之间是否存在真正的控制元件，可以将扫描该 35 bp 区域的 10 bp 置换突变引入 −50～+55 启动子片段。如果这些突变中有 1 个或 2 个再现了启动子活性的两倍损失，那么该区域中可能存在相关的控制元件。利用更小的置换突变体可以对假定的元件进行更精确的定位，然后具体分析其作用机制。如果利用置换突变体没有观察到这样的效应，说明缺失分析中观察到的效应可能与启动子活性无关。

关于缺失分析的数据要注意的第二点是，为一些突变体确定的标准差相当大。存在相当大偏差的原因是分析所用的数据来自以不同 DNA 制剂在不同时间中进行的多个实验。如果数据来源于用相同 DNA 制剂在同一天中所做的多个独立的转染实验，标准差就会低得多。然而，每个突变的效应在某种程度上随 DNA 制剂的不同和实验时间（天与天之间）的不同而有所差别，这可能归因于不同 DNA 制剂中有不同浓度的杂质污染，或细胞的生长状态、健康状况及转染效率的差异。虽然这些差异的存在会带来了较大的标准差，但较大的标准差能使数据结果更准确地呈现出来。如果标准差来源于使用唯一 DNA 制剂或只在一天中进行的实验，得到的数据可能会不够准确，进而可能会引起误导。

最后，每个突变体的数据以诱导后野生型启动子活性的百分比来表示（如将诱导的野生型启动子活性设为 100%，确定每个突变体诱导后相对于野生型的活性），而不以 LPS 的活化倍数来表示，因为后一数值可能依赖于非诱导信号的有效性。由于非诱导信号与背景相当接近，故对其有效性没什么信心，下文将进一步评论有关可诱导性的证据。

更为相关的缺失似乎是那些延伸超过 −150 的缺失。如图 6.1 所示，−143～−135 之间的缺失将启动子活性降到野生型启动子（即 −350～+55 片段）活性的 20%，−126～−123 之间的缺失将启动子活性降到野生型启动子活性的 10%，−102～−97之间的缺失将启动子活性降到接近背景。这些结果表明重要元件可能存在于 −143～−97 之间。此外，同等重要的元件可能存在于 −97～+55 之间。因为缺失至 −97 时，启动子活性降至了背景水平，但是这些数据没有提供 −97 下游存在重要元件的信息。还需要利用置换突变体或 3′ 端缺失突变体确定该下游区域中是否存在重要的控制元件。

对于 *Il12b* 的研究来说，没有从启动子的 3′ 端来制备缺失突变体（即从 +55 向转录起始位点依次缺失序列）。这是因为 −150～+55 序列大小正适合通过置换突变体分析法进行解析。此外，因为启动子片段只包括 55 bp 的非翻译前导序列，为了不影响核心启动子元件，包括 TATA 框和潜在的起始位点序列，3′ 端只能缺失一小部分核苷

酸。因为这些核心启动子序列的缺失会影响基本转录机器在启动子上形成稳定起始前复合物的能力，所以难以对这一区域的缺失进行评价。因此，我们确定特定的置换突变将提供更多的信息。尽管如此，3′ 端的缺失体也能为此项研究提供额外的信息，并且可能对其他研究有利。

置换突变体分析

在－150～＋55 区引入特定的置换突变体，可以鉴定重要的控制元件（图 6.2 和图

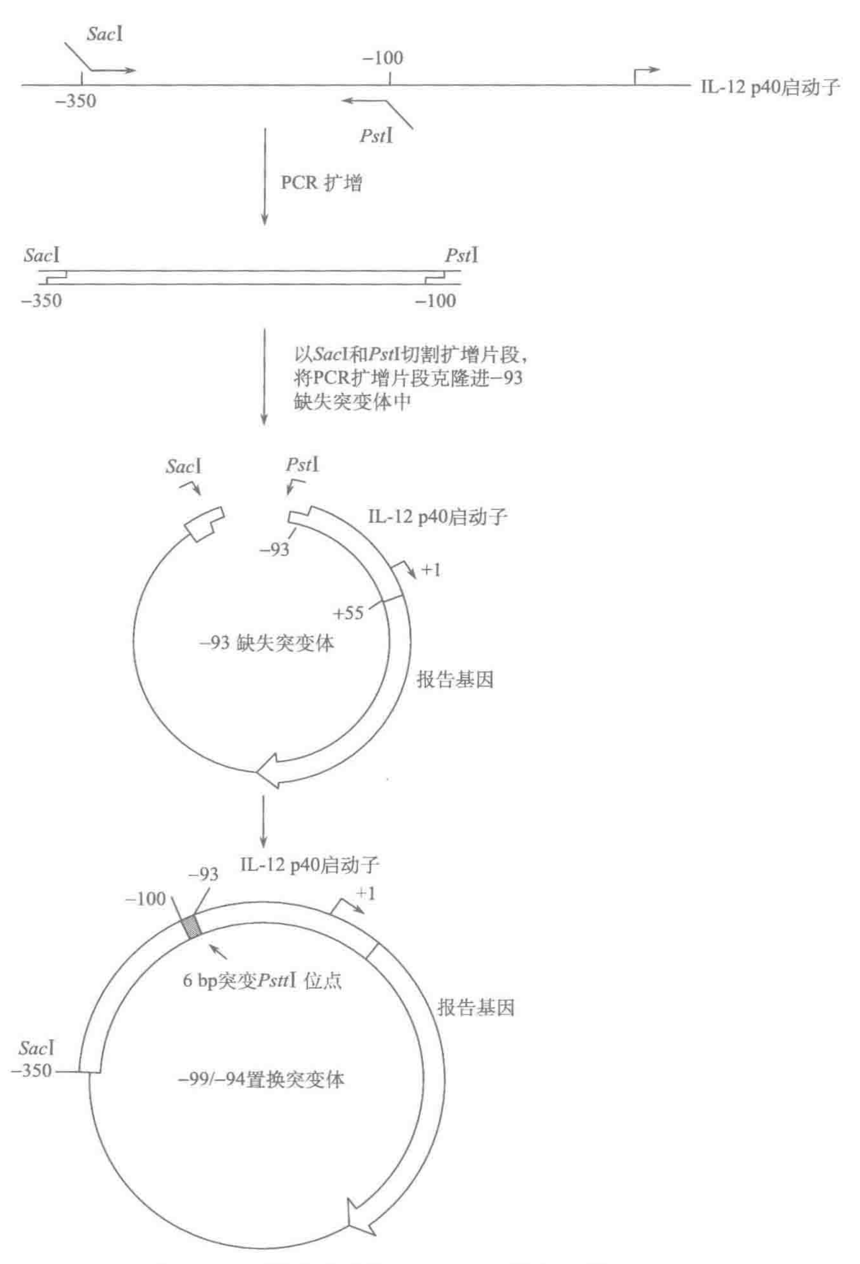


图 6.2 从缺失突变体系列产生置换突变体的策略。

6.3)。虽然－150～＋55 片段仅保留 40%的野生型启动子活性，但该活性仍是强烈可诱导的（未显示），因此，该区域很可能即使不包含参与可诱导转录的全部关键启动子元件，也包含参与可诱导转录的大多数关键启动子元件。引入－150～＋55 之间的大多数突变改变了 6 bp，虽然也有一些改变了 5 bp，还有一些改变了 10 bp（图 6.3）。

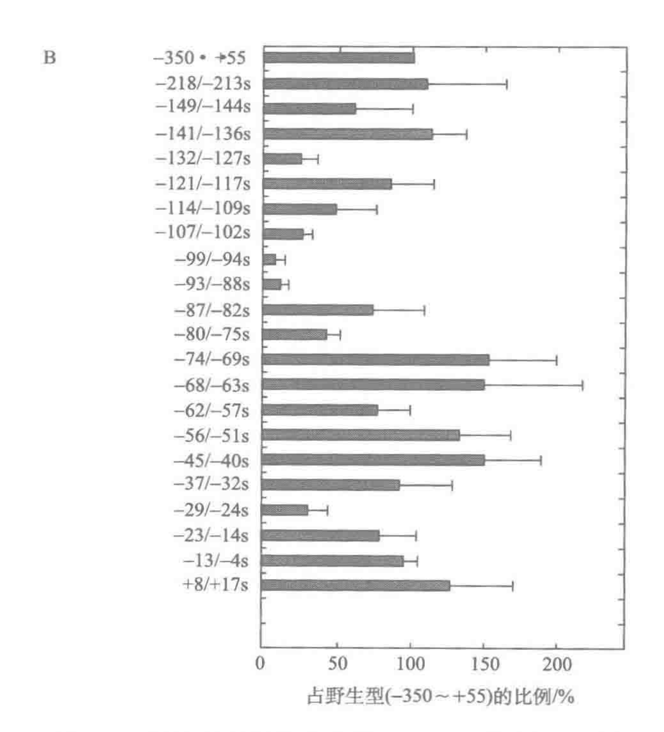


图 6.3 鼠 *Il12b* 启动子的置换突变体分析（经许可，改编自 Plevy et al. 1997，© American Society for Microbiology）。

为了构建大多数突变型质粒，用上述 5′ 端缺失构体作为起点。为了从缺失突变体产生置换突变体，先用 PCR 扩增启动子的远端部分，再将 PCR 产物与含有近端部分的适当缺失突变体融合，这样融合产生的质粒将在融合位点含有一个置换突变。例如，置换突变体－99/94 是从缺失突变体－93 处产生的（图 6.2）。该缺失突变体在启动子－93核苷酸的上游直接紧邻一个 *Pst*I 位点，*Pst*I 位点的上游又直接紧邻一个 *Sac*I 位点。为了产生置换突变体，用一个含有 *Sac* I 限制性位点的上游引物和一个含有 *Pst*I 位点的下游引物，通过 PCR 从全长启动子模板扩增－350～－100 之间的启动子序列。然

后将 PCR 产物插入经 *Sac*I/*Pst*I 切割的－93 缺失突变体质粒。在－350/＋55 启动子片段上下文序列中，*Pst*I 位点产生了一个－99 ～－94 的 6 bp 突变。因此，对于该项分析来说，缺失突变体充当许多置换突变体的克隆中间体。用于产生置换突变体的替代策略（如 Stratagene QuikChange 方法）也已同样获得了成功。

之所以使用 5～10 bp 的置换突变，有两点理由。第一，突变要足够小，以便重要的控制元件可以以合理的精度定位。第二，突变体最好要足够大，以保证所需的突变体的数量合理。通过使用 21 个 5～10 bp 的突变体，就能扫描－150～＋55 的几乎整个区域内的功能性元件。

图 6.3 所示的突变体系列中可以看到几个明显的缺口。4 bp 或更小的缺口可能无太大意义，因为大多数序列特异性 DNA 结合蛋白识别 6 bp 或更大的序列。因此，位于这些小缺口侧翼的突变应提供关于控制元件存在的信息。5 bp 或更大的缺口增加了漏掉重要控制元件的可能性。在最初的 *Il12b* 分析之后保留的大的缺口不是该分析所期望的，并且这些大缺口随后在其他突变的分析中被消除（J. Gemberling and S. Plevy，个人通讯）。

通过在瞬时转染分析法中分析置换突变体，可以鉴定重要的控制元件（图 6.3）。在此分析中，对启动子活性最重要的元件明显被－99/－94 突变体和－93/－88 突变体破坏了。这些突变体把启动子活性降至野生型启动子活性的 10%左右。另一个突变，即－132/－127，将启动子活性降至野生型启动子活性的 25%，表明它干扰了另一个重要元件。还有另外两个突变，即－107/－102 和－29/－24，将启动子活性降至野生型启动子活性的 25%左右，后一突变破坏了 TATA 框，而前一突变体接近两个严重的突变，表明它可能影响同一元件。

许多剩余的突变体都表现出介于野生型启动子活性 50%～150% 之间的启动子活性。这些微小的效应表明存在对启动子活性较不重要的元件，或是其活性与其他元件（见下文）的活性很大程度上冗余（或至少非协同性）的元件。此外，如上文对－350～－150 之间缺失突变体所做的讨论，这些微小的效应可能是由在启动子内特异性位点中外源序列的引入造成的。也就是说，突变的序列可能不含重要元件。值得注意的是，随后的研究揭示了激活蛋白－1（AP-1）、活化 T 细胞核因子（NFAT）及干扰素调节因子（IRF）家族成员（如 ICSBP）重要结合位点的存在，且这些重要结合位点都位于－88和 TATA 框之间（Zhu et al. 2001，2003［见图 6.7］）。在最初的分析中，这些元件由于冗余性被遗漏了，因为最初引入到这些元件之一中的置换突变体没有充分破坏蛋白质的结合和（或）因为这些元件对 *Il12*b 启动子活性的重要性对转染和细胞活化的条件是敏感的。

精细的置换突变体分析。

图 6.3 中的结果表明，在大约－99 和－88 之间存在一个关键的控制元件，而另一个关键的控制元件可能位于大约－132 和－127 之间。在这点上，可以认为突变体分析已经完成了，接下来应该用实验来鉴定与－99/－88 和－132/－127 两处的元件结合的蛋白质。这一说法有一定的正确性，然而，重要的是要解释为什么其他突变体的构建和分析可产生新的、有价值的信息。

5～10 bp 置换突变体获得的结果带来的问题是，这些结果提供了关于重要元件边界的相当不精确的信息。例如，－99/－88 元件内的重要核苷酸可能从－101 延伸至－88或更远，因为－107/－102 突变将启动子活性降低了 4 倍。另外，－132/－127 元件内的重要核苷酸可能从－135 延伸至－122，因为对启动子活性没有显著影响的侧翼突变是－121/－117 和－141/－136 突变。

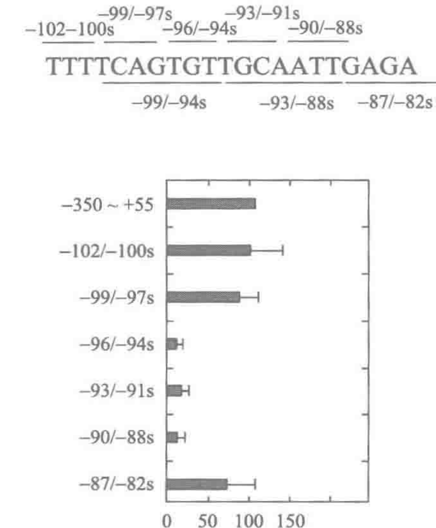


图 6.4　鼠 *Il12b* p40 启动子的精细的置换突变体分析（经许可，改编自 Plevy et al. 1997，© American Society for Microbiology）。

在解释界定元件边界的理由之前，我们先描述一下对－99/－88 元件的研究策略和所得结果。其策略是产生一系列突变体，每个突变体改变 3 个相邻的碱基对（图 6.4）。对 3 bp 突变体系列的分析结果揭示出有 3 个突变体大大降低了启动子活性。相比之下，侧翼突变对启动子活性没有显著的影响。这些发现表明关键核苷酸最少跨越 5 bp（即－94～－90），而最多跨越 9 bp（－6～－88)。这与一个蛋白质结合位点的存在最为一致。事实上，TRANSFAC 数据库的搜索揭示了关键的 9 bp 序列代表了 CCAAT/增强子结合蛋白（C/EBP）家族成员（Wedel and Ziegler-Heitbrock 1995）的结合位点。随后的 DNA 结合研究支持了 C/EBP 蛋白与关键元件发生功能性相互作用这一假设（Plevy et al. 1997；Bradleyetal 2003)。用一系列 3 bp 突变分析鉴定出的－130 处的功能性重要核苷酸与任何已知蛋白（未出示）的一致性结合位点都不匹配。然而，对特定核苷酸序列的仔细检查发现它代表一个 NF-κB 蛋白的非一致性结合位点（Murphy et al. 1995；Plevy et al. 1997；Sanjabi et al. 2005)。

我们现在回过头来看一下这个问题：为什么值得对每个元件的边界进行精确定位？答案是：因为对元件边界的了解将帮助我们确定每个元件是否与一个关键的序列特异性 DNA 结合蛋白相互作用，或者每个元件是否代表一个混合元件，该混合元件含有两个或三个蛋白质的相邻的结合位点。这一点很重要，因为已经发现混合元件以相当高的频率存在于启动子和远程控制区中。混合元件通常含有两个或三个蛋白质相邻的结合位点，这些蛋白质协同性地结合，和（或）以协同的方式发挥作用。如果发现一个重要元件跨越了 5～10 bp，那么这个元件最有可能只与一个蛋白质相互作用（或者是含有同时发生的两个蛋白质的结合位点的混合元件，如 NF-κB/HMGA 混合位点）（Thanos and Maniatis 1992)。另一方面，如果一个元件跨越 15～20 bp 或更长，那就更有可能是一个混合元件。通过分析几个额外的突变就能获得对此问题的见解，而且付出这种努力是有价值的。没有这个信息，我们就要依赖于结合研究洞悉与每个元件发生功能性相互作用的蛋白质的数目。考虑到难以确定哪个结合蛋白与控制元件是相关的（见上文和

第5章)，所以仅依赖于结合研究确定一个混合元件是否存在是不可取的。在对－99/－88的分析中，3 bp突变体分析结果表明该元件最大由9 bp组成。当发现9 bp是C/EBP蛋白的近一致位点时，就可以相当有信心地进行蛋白质研究了。如果尚未确定元件的边界，就不能确定元件是否还含有其他蛋白质的结合位点的可能性。

3 bp突变体分析的另一个好处值得注意：在鉴定了结合活性之后，就可以评估蛋白质与不同突变体相结合的能力。蛋白质结合所需的核苷酸和转染实验中控制元件发挥作用所需的核苷酸之间的密切相关性，提供了可支持该蛋白质功能相关性的一项重要数据（见第5章)。在这种情况下，重组C/EBP蛋白的结合所需的核苷酸正是启动子发挥功能所需的核苷酸，这支持了有关C/EBP蛋白负责该元件功能的假设。

关于精细的突变体分析应讨论的最后一个问题是：为什么用3 bp突变体而不用单碱基对突变体？因为用单碱基对置换将需要三倍之多的突变体才能扫描该重要区域。应当承认，用单碱基对突变体能更精确地界定元件的边界，但对许多研究来说，这点好处可能不会胜过需要额外耗费的精力。此外，许多蛋白质都能忍受其结合位点内某些位置上的单碱基对变化，随之也只丧失一点活性。而且，如果使用单碱基对突变体，结果将取决于引入的特定核苷酸，因为结合蛋白经常在特定位置上对某些碱基置换的忍受能力比对其他碱基置换的忍受能力强，而DNA结合蛋白忍受特定3 bp突变的可能性就低得多了。

用于置换突变体的核苷酸的选择。

遗憾的是，当构建置换突变体时，没有针对核苷酸选择的万无一失的策略。总是存在这样的可能性，即一种突变将产生另一蛋白质的一个偶然的结合位点，并且该蛋白质可能影响启动子的活性。新结合位点的产生可能得出不准确或误导性的数据。在5～10 bp置换突变体系列中，代替*Il12b*启动子序列插入限制性酶切位点。在这种情况下，限制性位点对诱变来说是必要的，因为从缺失突变体将产生置换突变体，使用的技术依赖于限制性位点的存在（见图6.2)。对于大多数诱变方法来说，限制性位点的插入很方便，因为它简化了对那些含有突变质粒的细菌菌落的筛选过程；通过限制酶切作图能够简单地分析DNA的小量制备物，而不用进行DNA测序。

对于3 bp突变体来说，因为只改变了3 bp，所以不能通过常规方法引入限制性位点。因为置换的部分小，所以要尽可能大幅度地改变序列，以增加重要的蛋白质-DNA接触被破坏的可能性。最后，每次置换突变都引入碱基颠换，同时改变碱基对。例如，C/G碱基对变为A/T、G/C变为T/A、A/T变为C/G、T/A变为G/C，这些置换可能在某一特定位点引入最彻底的结构改变。然而，这一做法没有考虑置换会导致产生另一种蛋白质结合位点的可能性。

为了确定是否产生了另一蛋白质的结合位点，可以通过搜索TRANSFAC这样的结合位点数据库来分析突变序列。如果突变序列与已知蛋白的一致序列相似，就应引入与之不同的序列。当然，这种分析只会基于目前的知识揭示已知蛋白的结合位点。因为没有决定性的方法来确保一个置换突变不引入其他结合位点，所以重要的是在启动子分析的随后步骤中要记住这种可能性。如果在分析的任何阶段，结果表明可能已经意外地引入了一种偶然性的结合位点，那么最好的行动方针是通过在相同的位置引入其他的置换突变来解决这种可能性。

可诱导性和细胞类型特异性

利用上面概括的诱变策略导致了两个控制元件的鉴定，这两个元件促成 LPS 诱导巨噬细胞中 *Il12b* 启动子的活性。该启动子是看诱导的，所以这两个元件都可能与直接调解诱导的蛋白质结合。或者，其中一个元件可能与一种组成型活性蛋白结合，该蛋白质对启动子的功能是必不可少的，但却不是诱导的直接调解因子。在后一种推测中，当组成型活性蛋白与诱导蛋白发生适当的生理性或功能性相互作用时，启动子的诱导则发生。因为启动子分析的主要目标是阐明启动子诱导的机制，所以最终有必要区分这两种类型的元件。当研究细胞类型特异性控制区时，也必须考虑类似的问题。在这些研究中，目的是区分两种控制元件，即结合细胞类型特异性蛋白质的控制元件和结合遍在性活性蛋白的控制元件。还要补充的重要一点是某些因子能够通过同一 DNA 模体对同一基因的转录抑制与转录激活都起作用，这更增加了了解每个模体的精确功能的挑战性。

很难确定哪个元件直接介导诱导，因为诱导后控制区的活性对任何控制元件内的突变都是敏感的，这里所说的任何控制元件也包括那些不介导诱导而仅为控制区活性所需的控制元件。因为诱导后一个元件的重要性不能提供有关可诱导性本身的重要信息，所以用于确定哪个元件是诱导的直接调解者的常见策略是依赖于诱导前各元件的重要性。理论上，与直接调解诱导的蛋白质结合的元件在非诱导细胞中不会参与基本启动子活性，并且将只有在诱导后才变得重要起来。相比之下，结合组成型活性蛋白的元件在非诱导和诱导细胞中同样重要。因此，通过比较野生型启动子的和每个启动子突变体的“诱导倍数”，就可以简单地确定每个元件的准确作用。换句话说，去除背景后，每个构体的诱导后报告活性除以同一构体的非诱导活性，就得出诱导倍数值。如果一个元件的突变引起诱导倍数值减小（即该元件在诱导细胞中比在非诱导细胞中更重要），则认为该元件是诱导作用的直接调解者。如果一个元件不影响诱导倍数值（即该元件在诱导细胞中和在非诱导细胞中同等重要），则该元件不可能介导诱导作用。

以上策略对于在诱导前产生相当大活性的部分启动子来说，能提供一些信息。但对于非诱导活性尚未明显高于背景的启动子来说则没有用。例如，*Il12b* 启动子的活性在诱导前只稍高于背景。虽然野生型启动子的非诱导活性在统计上是显著的，但重要的控制元件内的突变将非诱导信号降到不显著高于背景的水平，这一相当常见的情况使得对诱导倍数值的分析变得毫无意义。

作为一个特定的例子，我们考虑一下 NF-κB（Rel）位点内的突变，该突变让诱导的启动子活性降至野生型启动子活性的 25%（见图 6.5）。该突变也将非诱导信号值降到稍高于背景的水平。因为扣除背景后启动子突变体的非诱导信号几乎是 0，诱导倍数却依然很高。如果把这些值作为确定一个元件是否介导可诱导性的唯一标准，可能会推断 NF-κB 位点对诱导不重要，尽管有大量关于 NF-κB 蛋白在巨噬细胞基因诱导期间起着关键的、直接作用的证据。

C/EBP 位点内的突变结果提供了另一个实例。随着突变，诱导的启动子活性降至野生型启动子活性的 10%以下（图 6.5）。非诱导启动子活性也降低了，但如 NF-κB 突变体一样，非诱导信号仍然要比背景略高一点。计算得到的 C/EBP 突变体的诱导倍数比 NF-κB 突变体的诱导倍数要低得多。对于这些数据的一种解释是：C/EBP 是诱导作

用的关键位点，而 NF-κB 位点的重要性要低得多。虽然这种说法似乎符合逻辑，但实际上是对数据的曲解，因为诱导倍数值取决于非诱导信号的精确性。因为 NF-κB 突变体和 C/EBP 突变体及野生型启动子的非诱导信号都与背景接近，所以难以确定这些诱导倍数的准确性。换句话说，非诱导信号中非常小的变化也能对诱导倍数产生显著的、但备受争议的影响。

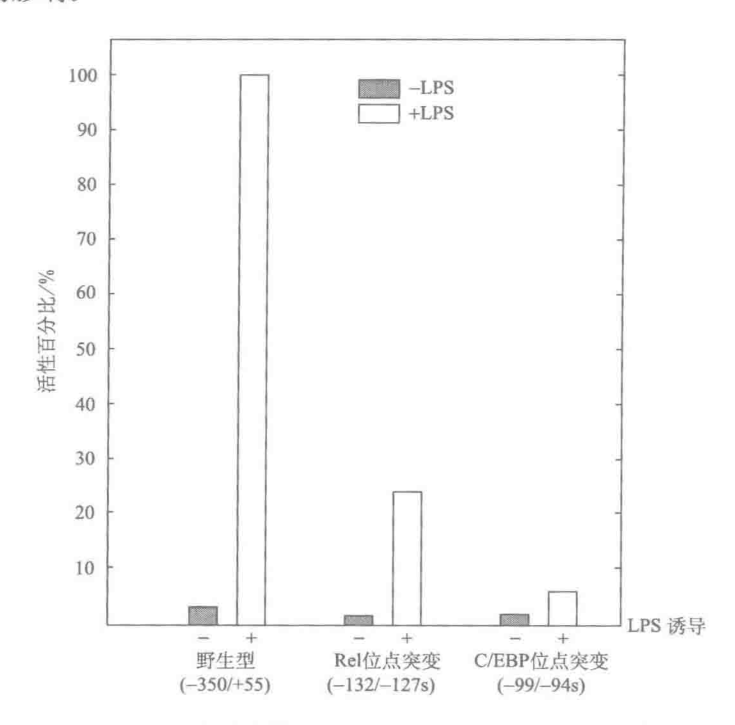


图 6.5 *Il12b* 突变体对脂多糖引起的启动子诱导的影响。

不管获得的结果如何，启动子突变体的活性很少实质性地洞悉哪些元件直接地负责可诱导性（或细胞类型特异性）。为了解决这个问题，必须鉴定相关的结合蛋白并表征其特征。如果在细胞诱导期间相关的转录因子的丰度增加，那么该转录因子可能直接地对启动子的诱导起作用。可以用其他方法更严格地验证这一假设，其中有些方法会在第 8 章进行讨论。如果在细胞诱导期间转录因子的丰度没有增加，那么该转录因子仍然可能在诱导过程中起直接作用，因为很可能该转录因子需要翻译后修饰来改变其活性。最后需要对每个结合位点和转录因子进行认真分析以确定哪些结合位点和转录因子在诱导或细胞类型特异性中起直接的作用。

转录起始位点的确证

理想情况下，当用瞬时转染分析启动子时，应确定转录起始位点，以确证转录在正确的位置上起始（见第 3 章和第 4 章）。还应确定突变启动子的转录起始位点，以确证突变没有改变起始位点或引起隐秘起始位点的诱导。不用顾虑启动子诱变时起始位点发生较小的改变，但起始位点发生较大的改变可能表明突变已不是简单地破坏了控制元件。相反，突变可能已经改变了启动子的整体结构和调控。例如，突变可能会引起启动

子中突变核苷酸内或其他位置隐秘的类 TATA 序列的活化。如果隐秘的类 TATA 序列被活化了，它对调控元件的应答可能与真正的核心启动子不同。

虽然许多细胞系的瞬时转染效率很低，以至于不能通过引物延伸或核糖核酸酶（RNase）保护来对起始位点作图，应当是用 5′ RACE（cDNA 末端快速扩增）来洞悉正在使用的起始位点（见第 3 章）。作为灵敏的 5′ RACE 方法的替代方法，可用启动子报告质粒制备稳定转染体。因为经筛选的细胞系中的每一个细胞将包含一个整合的报告质粒，所以报告转录本有足够的丰度用于通过引物延伸和核糖核酸酶保护法对起始位点作图。此策略被用于 *Il12b* 启动子的分析，并且引物延伸的结果确证了主要起始位点在预计的位置上。这个结果提供了一些信心，使人相信在瞬时转染质粒中转录起始位点也可能在正确的位置上。

因为在瞬时转染之后难以对转录起始位点进行定位，所以在一般的突变体分析过程中通常不进行瞬时转染实验。对于解析核心启动子区（即 TATA 区和 Inr 区）的有关研究来说，缺少起始位点位置的信息可能是成问题的。然而，对其他大多数研究来说，起始位点的分析不是必不可少的。不过，仍要相当谨慎地进行研究，并时刻注意起始位点还未确证这个事实。

分析方法的选择

诱变策略的一个关键的局限性是只能鉴定那些在所用的功能分析中重要的元件。在用瞬时转染分析法对 *Il12b* 启动子进行分析时，C/EBP 位点是必不可少的，而 NF-κB 位点只起中等程度的作用。相比之下，当在稳定转染分析中对这两个位点的突变进行检验时，这两个位点对启动子活性来说都是绝对必不可少的（Plevy et al. 1997；Sanjabi et al. 2005），推断瞬时转染细胞中质粒的高拷贝数或瞬时转染质粒的附加性降低了 NF-κB 位点的重要性（见第 4 章）。因此，若说利用稳定转染分析法重复整个突变体分析鉴定到了在瞬时转染分析中相对不重要的其他的必不可少的控制元件，这也并不令人意外。事实上，已经报道过几个在稳定转染中起作用而在瞬时转染中不起作用的控制元件（见第 4 章）。

瞬时转染和稳定转染分析都可能会漏掉一些对于内源性基因的转录来说重要的控制元件，这是因为稳定转染质粒不能整合进那些与内源性基因相同的染色质结构。所以说，可能会漏掉一部分在基因活化或失活期间对染色质重塑来说重要的控制元件。

免疫球蛋白（Ig）μ 内含增强子的分析为上述情况提供了一个实例。在转基因鼠中，邻近的基质附着区（MAR）强有力地刺激了 Ig μ 增强子的活性（Forrester et al. 1994，1999；Fernandez et al. 1998）。然而，在瞬时转染和稳定转染分析中，MAR 对增强子活性都没有影响。因此，如果只利用转染分析来解析 Ig μ 增强子，就会忽视 MAR 的重要性。有趣的是，Forrester 和 Grosschedl 发展了一种改进的稳定转染分析，该分析恢复了对 MAR 的需要（Fernandez et al. 1998；Forrester et al. 1999）。为了进行此项分析，在转染和药物筛选之前，对报告质粒 DNA 进行体外甲基化。来自稳定整合并预先甲基化的质粒的报告基因活性需要 MAR 的存在。显然，预甲基化使转染质粒整合进可及性较差的染色质，这将引起转录激活对 MAR 的需要。

基于以上结果，理想的情况是，利用转基因鼠分析法或稳定转染分析法对一个新的

控制区进行突变体分析。遗憾的是，不管用这些方法中的哪一种，综合性突变体分析所要求的时间和资源对大多数研究来说都是不现实的。由于这种原因，通常有必要从瞬时转染分析法开始鉴定综合性突变体分析中所需的控制元件，之后，当分析达到更高级的阶段时，再进行更复杂的分析。

控制元件的冗余性

关于综合性突变体分析的最后一个警告是，此法可能无法鉴定那些活性与控制区内其他控制元件的活性相冗余（或至少不是强烈的协同性）的控制元件。Ig μ 增强子提供了一个控制区内冗余性的一个典型例子。早期的转染研究发现，没有任何置换能把活性降至两倍以上（如见 Lenardo et al. 1987），转基因鼠分析也得出相似的结果（Jenuwein and Grosschedl 1991；Annweiler et al. 1992），这表明明显的冗余性不是转染分析的实验假象。任何突变都不能强有力地降低增强子的活性，这给对 Ig μ 增强子作用机制的分析造成相当大的困难。

然而，Sen 及其同事发展了一种策略，以用于克服冗余性问题和解析冗余性的分子基础（图 6.6）。他们首先构建了缺失突变体，以鉴定优先在 B 细胞中支持增强子功能的最小增强子片段（Nelsen et al. 1990，1993）。正如根据观察到的冗余性所预期的，可以缺失几个控制元件而几乎不产生后果。然后，在最小增强子片段内引入置换突变，发现剩下的控制元件对于增强子功能来说是绝对必不可少的（Nelsen et al. 1990，1993）。对必不可少的控制元件的进一步分析鉴定了可能与这些控制元件发生功能性相互作用的蛋白质（Nelsen et al. 1990，1993）。利用最小增强子片段已经解析了这些蛋白质彼此之间协同作用的分子机制（Erman and Sen 1996，1997；Rao et al. 1997；Erman et al. 1998）。最初的研究使用了只包含三个控制元件，即 μA、μE3 和 μB 的最小增强子片段（图 6.6）。尽管有了这个小片段，还必须在报告质粒中融入二聚体以检测活性。在鉴定和表征完这一片段内的三个控制元件以后，再使用一个在单拷贝时就能产生大量的增强子活性的较大片段（Dang et al. 1998b）。因为该较大的片段仍缺少冗余的元件，所以大多数元件对增强子活性来说仍是必不可少的。

为了了解冗余性的分子基础，Sen 及其同事系统突变了全长增强子，以鉴定与 μE3 相冗余的控制元件（图 6.6），目的是在增强子含有突变体 μE3 位点时，鉴定对功能关键的元件。结果发现，赋予冗余性的元件是一个之前未描述过的、与 IRF 蛋白结合的增强子元件（Dang et al. 1998a）。因为 IRF 元件主要与 μE3 元件相冗余，所以它似乎对增强子功能来说一样重要，即使是它在 μE3 元件被首次报道后 13 年才被发现。

值得注意的是，μE3 元件最初是通过体内和体外蛋白质-DNA 相互作用研究被鉴定出来的。相比之下，IRF 元件的发现需要系统的突变体分析，这种分析是全面的，足以解决冗余性问题。现在，这种系统分析可被进一步拓展，以用于确定增强子内的其他元件是否对冗余性起作用。冗余性的生物学基础依然未知。一种假设是，在发育的不同阶段，冗余性使增强子被不同组合的因子激活，以致无论何时增强子被激活，都不需要存在一套相同的特定因子。或者，冗余性可能确保在极限转录因子浓度的细胞核环境中增强子的功能。

Sen 及其同事所用的策略很可能对于分析其他控制区内的冗余性（或非协同性）元

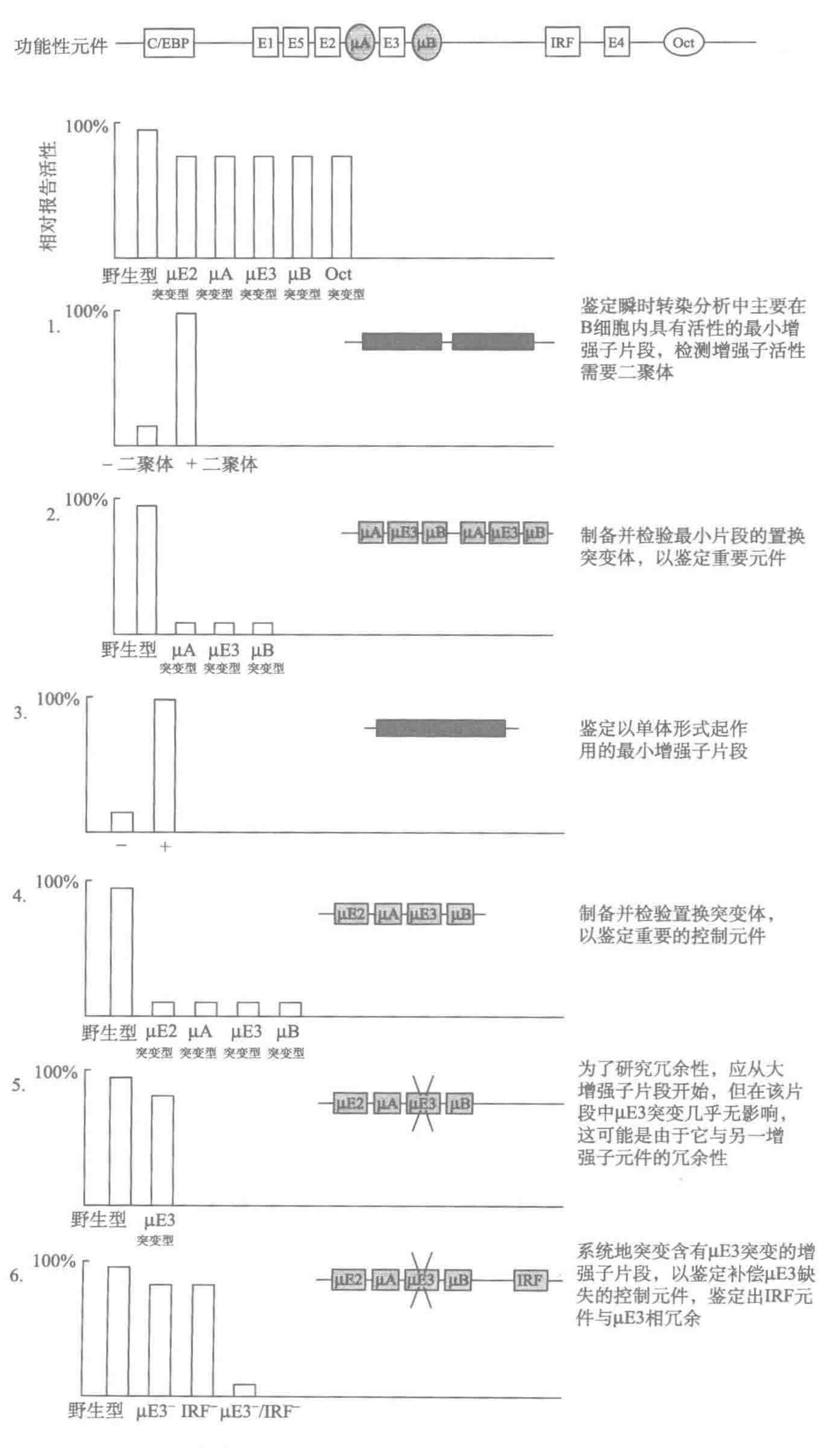


图 6.6　用于解析在控制元件中显示很大冗余性的增强子的策略。

件是有用的。某些冗余性可能具有生物学意义，而其他的可能与分析所用的分析方法有关。Ig μ 增强子似乎是生物学相关冗余性的一个例子，因为在转基因鼠分析及瞬时转染和稳定转染分析中都观察到 Ig μ 增强子的冗余性。

来自综合性突变体分析与系统发生分析比较的深刻见解

通过与从系统发生分析及结合位点数据库分析中获得的信息的比较，现在我们讨论一下上面所描述的 *Il12b* 启动子的综合性突变体分析的成功水平，并以此结束本章。图 6.7 展示了通过加利福尼亚大学 Santa Cruz 基因组浏览器产生 *Il12b* 启动子的系统发生分析（也见第 5 章）。在本章所讲述的综合性突变体研究中，以及在其他发表的研究中，原创性鉴定的 DNA 元件的位置被描画在序列的上面（Zhu et al. 2001，2003；Murphy et al. 2005；Saniabi et al. 2005；Plevy et al. 2007）。有趣的是，只有 TATA 框和 C/EBP 位点能轻易地在基本 TRANSFAC 数据库搜索中检测到，其他位点与现在已知的能与这些位点相结合的蛋白质的报道过的一致序列不同。在以降低的严紧性进行的 TRANSFAC 分析中，鉴定出几个额外的位点（未出示），但在严紧性降低的分析中，这些功能性重要位点并没有显示出与预测的其他大量转录因子的非一致序列位点有关。例如，TRANSFAC 分析预测出了位于进化中不保守的位置，以及在瞬时转染分析中对启动子活性不重要的位置上的 NF-κB 结合位点，但 TRANSFAC 分析却漏掉了能与 NF-κB 结合的功能性重要位点。

值得注意的是，系统发生分析揭示了通过哺乳动物的进化，所有功能性重要转录因子的结合位点总是高度保守的。因此，如果在进行突变体研究之前，已经获得了如此大量物种的基因组序列，系统发生分析将能成功地预测出这些元件具有功能重要性。有趣的是，系统发生分析还揭示了几个其他同样高度保守的 DNA 模体（图 6.7 中编号 1～8 的水平线），即使通过诱变破坏这些模体，也对瞬时转染分析中的启动子活性毫无影响（见图 6.1 和图 6.3）（Plevy et al. 1997）。

以上观察引起了一个关键的问题：这些在瞬时转染分析中对 *Il12b* 启动子活性不重要的高度保守的元件究竟对启动子的活性重要吗？一种可能性是这些元件在不同类型的细胞中，或对不同刺激的应答中，促成启动子活性。或者，这些元件可能只在更自然的染色质环境中对转录是重要的，因为在那样的环境中它们可能促进了启动子和远距离增强子之间的相互作用，或者促成了在转染分析中对转录不是必需的核小体重塑事件。最后一种可能性是，这些元件可能在转染分析中（并且也在内源性基因座上）的确是重要的，但它们的功能可能不明显，因为它们与其他启动子元件冗余性地发挥作用。为了辨别这些可能性，有必要仔细评价这些高度保守的元件在其他情况下的功能，特别是在天然染色质环境中通过诱变考察它们的功能。

基于这个比较，我们断定阐明调控任何哺乳动物基因的机制需要各种各样的实验方法。通过系统发生分析、结合位点数据库分析，以及 ChIP 实验来评价预测的蛋白质-DNA 相互作用，可得到基因调控机制的初步认识。使用转染分析法所进行的综合性突变体分析很可能揭示至少几个在转录调控中起特别重要作用的关键元件，并可让研究人

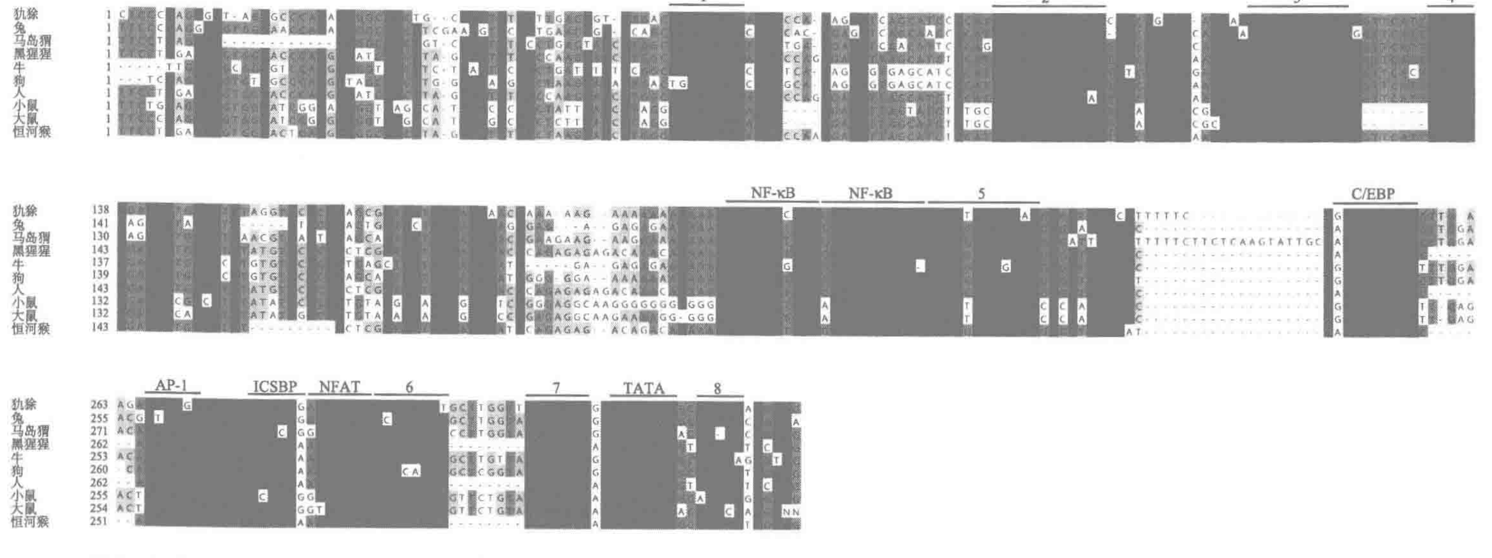


图 6.7 *Il12b* 启动子的综合性突变体分析和系统发生分析的比较所提供的信息。系统发生分析利用了加利福尼亚大学的 Santa Cruz 基因组浏览器(http://genome.ucsd.edu;Kent et al. 2002)。

员有信心启动对结合这些元件的因子的更高级的研究。然而，要充分了解基因的调控，必须发展用于探索许多在进化中保守但在标准转染分析中不起重要作用的其他DNA模体的功能作用的策略。而接近这一目标的第一步将可能是在自然染色质环境中对保守元件进行综合性突变体研究。

参考文献

Annweiler A. , Muller U. , and Wirth T. 1992. Functional analysis of defined mutations in the immunoglobulin heavy-chain enhancer in transgenic mice. *Nucleic Acids Res*. 20: 1503-1509.

Bradley M. N. , Zhou L. , and Smale S. T. 2003. C/EBPβ pregulation in lipopolysaccharide-stimulated macrophages. *Mol Cell Biol*. 23: 4841-4858.

Dang W. , Nikolajczyk B. S. , and Sen R. 1998a. Exploring functional redundancy in the immunoglobulin μ heavy-chain gene enhancer. *Mol. Cell. Biol*. 18: 6870-6878.

Dang W, Sun X. H. , and Sen R. 1998b. ETS-mediated cooperation between basic helix-loop-helix motifs of the immunoglobulin μ heavy-chain gene enhancer. *Mol. Cell. Biol*. 18: 1477-1488.

Erman B. , and Sen R. 1996. Context dependent transactivation domains activate the immunoglobulin μ heavy-chain gene enhancer. *EMBO J*. 15: 4565-4575.

Erman B. , Cortes M. , Nikolajczyk B. S. , Speck N. A. , and Sen R. 1998. ETS-core binding factor: A common composite motif in antigen receptor gene enhancers. *Mol. Cell. Biol*. 18: 1322-1330.

Fernandez L. A. , Winkler M. , Forrester W. , Jenuwein T. , and Grosschedl R. 1998. Nuclear matrix attachment regions confer long-range function upon the immunoglobulin μ enhancer. *Cold Spring Harbor Symp. Quant. Biol*. 63: 151-524.

Forrester W. C. , Fernandez L. A. , and Grosschedl R. 1999. Nuclear matrix attachment regions antagonize methylation-dependent repression of long-range enhancer-promoter interactions. *Genes Dev*. 13: 3003-3014.

Forrester W. C. , van Genderen C. , Jenuwein T, and Grosschedl R. 1994. Dependence of enhancer-mediated transcription of the immunoglobulin μ gene on nuclear matrix attachment regions. *Science* 265: 1221-1225.

Jenuwein T. and Grosschedl R. 1991. Complex pattern of immunoglobulin μ gene expression in normal and transgenic mice: Nonoverlapping regulatory sequences govern distinct tissue specificities. *Genes Dev*. 5: 932-943.

Kent W. J. , Sugnet C. W. , Furey T. S. , Roskin K. M. , Pringle T. H. , Zahler A. M. , and Haussler D. 2002. The Human Genome Browser at UCSC. *Genome Res*. 12: 996-1006.

Lenardo M. , Pierce J. W. , and Baltimore D. 1987. Protein-binding sites in Ig gene enhancers determine transcriptional activity and inducibility. *Science* 236: 1573-1577.

Ma X. , Chow J. M. , Gri G. , Carra G. , Gerosa E. , Wolf S. E. , Dzialo R. , and Trinchieri G. 1996. The interleukin 12 p40 gene promoter is primed by interferon γ in monocytic cells. *J. Exp. Med*. 183: 147-157.

Murphy T. L. , Cleveland M. G. , Kulesza P. , Magram J. , and Murphy K. M. 1995. Regulation of interleukin 12 p40 expression through an NF-κB half-site. *Mol. Cell. Biol*. . 15: 5258-5267.

Nelsen B. , Kadesch T. , and Sen R. 1990. Cornplex regulation of the immunoglobulin μ heavy-chain gene enhancer: μB, a new determinant of enhancer function. *Mol. Cell. Biol*. 10: 3145-3154.

Nelsen B. , Tian G. , Erman B. , Gregoire J. , Maki R. , Graves B. , and Sen R. 1993. Regulation of lymphoid-specific immunoglobulin μ heavy chain gene enhancer by ETS-domain proteins. *Science* 261: 82-86.

Nikolajczyk B. S. , Nelsen B. , and Sen R. 1996. Precise alignment of sites required for μ enhancer activation In B cells. *Mol. Cell. Biol*. 16: 4544-4554.

Nikolajczyk B. S, Cortes M. , Feinman R. , and Sen R. 1997. Combinatorial determinants of tissue-specific transcription in B cells and macrophages. *Mol. Cell. Biol*. 17: 3527-3535 .

Plevy S. E. , Gemberling J. H. , Hsu S. , DornerA J. , and Smale S. T. 1997. Multiple control elements mediate acti-

vation of the murine and human interieukin 12 p40 promoters: Evidence of functional synergy between C/EBP and Rel proteins. *Mol. Cell. Biol.* 17: 4572-4588.

Rao E., Dang W., Tian G., and Sen R. 1997. A three-protein-DNA complex on a B cell-specific domain of the immunoglobulin μ heavy chain gene enhancer. *J. Biol. Chem.* 272: 6722-6732.

Saniabi S., willlams K. J., Saccani S., Zhou L., Hoffnlann A., Gerondakis S., Natoli G., and Smale S. T. 2005. A c-Rel subdomain responsible for enhanced DNA-binding affinity and selective gene activation. *Genes Dev.* 19: 2138-2151.

Thanos D. and Maniatis T. 1992. The high mobility group protein HMG I (Y) is required for NF-κB-dependent virus induction of the human IFN-β gene. *Cell* 71: 777-759.

Wedel A. and Ziegler-Heitbrock H. W. 1995. The C/EBP family of transcription factors. *Immunobiology* 193: 171-185.

Zhu C, Gagnidze K., Gemberling J. H., and Plevy S. E. 2001. Characterization of an activation protein-1-binding site in the murine interleukin-12 p40 promoter. Demonstration of novel functional elements by a reductionist approach. *J. Biol. Chem.* 276: 18519-18528.

Zhu C., Rao K., Xiong H., Gagnidze K., Li F., Horvath C., and Plevy S. 2003. Activation of the murine interleukin-12 p40 promoter by functional interactions between NFAT and ICSBP. *J. Biol. Chem.* 278: 39372-39382.

网络资源

http://genome.ucsc.edu：对于鉴定和评价真核基因组中保守的非编码区，加利福尼亚大学圣克鲁斯基因组浏览器（The University of California，Santa Cruz Genome Browser）是有用的工具。

7 鉴定 DNA 结合蛋白及其基因

要点

- 鉴定 DNA 结合蛋白最常用的方法是电泳迁移率变动分析法和 DNase I 足迹法，每种方法都有不同的优点和局限性。
- 数据库方法有助于鉴定与 DNA 元件结合的蛋白质。
- 为了鉴定编码 EMSA 或 DNase I 足迹法检测到的蛋白质的基因，通常需要提纯该蛋白质。
- 也可以考虑使用其他方法鉴定编码 DNA 结合蛋白的基因，包括单杂交筛选、体外表达文库筛选和哺乳动物表达克隆。
- 由于更先进的质谱技术的出现，鉴定能与目的 DNA 模体结合的蛋白质可能变得更加容易。

引言和概述

上一章推荐系统发生分析作为一种鉴定目的控制区内潜在的重要 DNA 序列模体的首选方法。配套的突变体分析可提供证据，表明至少一部分保守元件在功能上对控制区的活性是重要的，同时也可能发现被系统发生分析所遗漏的功能元件。利用共有识别序列的数据库（如 TRANSFAC 或 JASPAR）评价目的 DNA 序列模体，可以提示候选 DNA 结合蛋白，同时可利用电泳迁移率变动分析（EMSA）和染色质免疫沉淀（ChIP）分析评价预测的蛋白质是否在体外和活细胞中分别与确定的 DNA 模体结合。

如果这个基本策略支持一个或多个已知的因子参与了目的基因的调控的想法，那么研究人员可以直接阅读第 8 章。第 8 章讨论了一个具有挑战性的目标，即确定候选因子是否的确是目的基因的重要调控者。然而，如果这个简单的方法并没有揭示一个可行的候选因子，或者研究人员想对结合保守或功能 DNA 模体的蛋白质进行更全面的分析，那么本章中讨论的方法可能是必需的。

不像本书中讨论的绝大多数程序，所推荐的用于实验鉴定与目的 DNA 模体相互作用的蛋白质的策略，在过去的数年中没有发生引人注目的变化。然而，由于重大的技术进步，正在探索可能很快取代传统方法的新方法。为了解释这些未来可能发生的变化，我们注意到，我们经常鼓励研究人员发展生化分析方法，如 EMSA 或 DNase I 足迹分析，以使用生化方法纯化蛋白质并鉴定其基因前与目的 DNA 模体可结合的蛋白质的检测。可以为这项建议列举多个原因。尤其是，蛋白质与 DNA 相互作用分析方法的最初发展使研究者在蛋白质纯化期间就能监测结合活性，并且使研究者在质谱分析前增强信心，相信纯化的蛋白质就是所需的目的蛋白（图 7.1）。

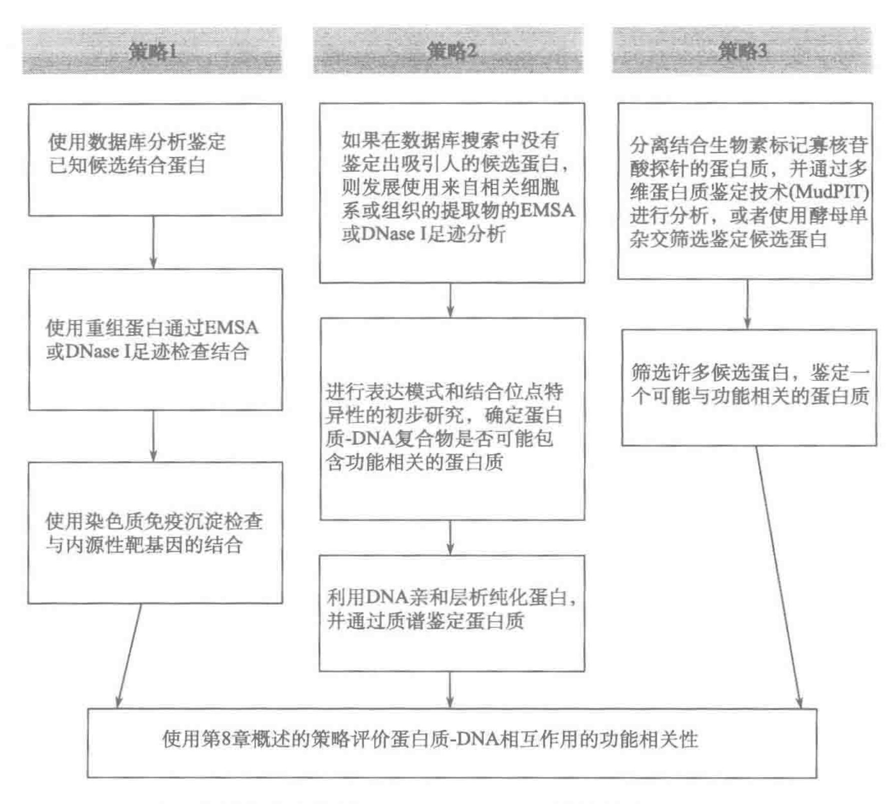


图 7.1　推荐用于表征与确定的控制元件结合的蛋白质的策略。

在不久的将来，质谱技术的巨大发展（Washburn et al. 2001；Domon and Aebersold 2006；Cravatt et al. 2007；Fournier et al. 2007），结合基因组序列信息的获得，不仅可能消除对初始发展蛋白质-DNA 相互作用分析方法的需要，而且可能消除对蛋白质纯化的需要。相反，很可能使用相对简单的捕获（pulldown）实验结合先进的质谱程序，如多维蛋白质鉴定技术（MudPIT）（Washburn et al. 2001；Fournier et al. 2007），快速鉴定细胞提取物中可与序列模体结合的所有蛋白质。其他不依赖于质谱的技术也开始出现，包括文库的发展，这种文库中包含表达为谷胱甘肽-*S*-转移酶（GST）融合蛋白的大量 DNA 结合蛋白，以便快速筛选出能与目的模体结合的蛋白质（Hazbun and Fields 2002）。这些进展可能极大地减少了鉴定 DNA 结合蛋白所需的时

间，将鉴定时间从一年或一年多减少到数周。剩下的障碍是能够获得的像 MudPIT 一样的技术非常有限，但这种限制很可能随时间的推移而减少。一个更具挑战性的障碍（见下文）是从 MudPIT 鉴定的蛋白质的集合中鉴定出序列特异性 DNA 结合蛋白，而不用花费大量的精力从以有限的特异性结合 DNA 的丰富的蛋白质中纯化分离序列特异性蛋白质。

预期在今后数年还会取得的另一重大进展是，基于质谱的用于分析从化学交联剂处理的细胞中分离的特定基因组座位的方法。这将使在自然染色质环境中与基因组座位相联系的蛋白质的全面的质谱分析成为可能。最近，Robert Kingston（个人通讯）及其同事发展了一种成功的策略，用于分析与多拷贝基因区相联系的蛋白质，未来的挑战是提高该技术的灵敏度，使其可用于单拷贝基因组座位的分析。在这一章中，我们首先讨论了如何利用共有序列数据库来鉴定可与目的 DNA 模体结合的候选蛋白。我们非常详尽地解释了为什么先进的质谱技术可能很快取代比较传统的蛋白质纯化策略，以用于鉴定 DNA 结合蛋白。之后，我们描述为细胞核提取物中与目的 DNA 模体可结合的蛋白质建立一种相互作用分析的传统方法，其中重点是 EMSA（凝胶迁移）和 DNase I 足迹分析。紧接着描述用于纯化由 EMSA 或者 DNase I 足迹检测到的 DNA 结合蛋白的传统 DNA 亲和层析程序。最后，我们描述不通过蛋白质纯化而鉴定编码目的 DNA 结合蛋白基因的策略，包括酵母单杂交筛选和表达克隆方法。由于篇幅所限，不包含这些方法的方案，但引用了恰当的文献出处。

重要的是要强调本章中讨论的每种策略都可能鉴定与目的 DNA 元件相结合的一个或多个蛋白质。然而，没有一种策略会一定导向在内源性座位环境中是控制元件的功能性调控者的蛋白质。不论 DNA 结合蛋白是怎样分离和鉴定的，它只不过是成为目的基因的候选调控者，每个候选者的功能相关性将需要用第 8 章所提供的策略进行评估。

鉴定 DNA 结合蛋白的实验策略

数据库方法

在第 6 章中，我们简单提及了数据库方法，并将其作为鉴定与目的 DNA 模体或控制区相互作用的转录因子的手段。最常用的两个数据库是 TRANSFAC（http://www.gene-regulation.com/pub/ databases.html）和 JASPAR（http://jaspar.genereg.net/），设计这两个数据库的目的是使研究人员确定 DNA 序列模体是否对应于之前所表征的蛋白质的潜在识别位点。不幸的是，这些数据库只包含那些结合位点经过严格的实验研究界定的蛋白质的高质量信息。此外，支持与 DNA 结合蛋白功能性相互作用的 DNA 序列范围的信息是不完整的，即使对于表征得最好的 DNA 结合蛋白也是如此。因此，在数据库搜索中，相关蛋白及其他家族成员被遗漏的概率仍然相当高。

用于结合位点数据库搜索的参数可以且应当是各种各样的。可以使用高严紧参数进行搜索，在这种搜索中只有与已报道的共有序列完全匹配的序列才会被揭示出来。然后，可以调整搜索参数，以搜索出与共同序列在一个或多个随机位置或特定位置错配的位点。我们强烈建议使用各种各样的参数进行分析，以鉴定尽可能多的候选蛋白。

如果目的 DNA 模体与一共有序列完美匹配，也只能说明预测的结合蛋白可能但不一定负责该模体的功能。第一种可能性是，在体外多个蛋白质都能与分离的位点结合，但只有其中一个负责体内元件的功能，并且该蛋白质可能还没有在数据库分析中被鉴定。第二种可能性是，共有序列未必完全准确，用于建立 TRANSFAC 和 JASPAR 等数据库的共有序列一般未被多个研究人员验证，而通常是由单一的研究报告的数据得出。第三种可能性是数据库中包含的共有序列可能是不完整的。例如，在数据库中有一些转录因子只有 4 个或 5 个碱基对的共有序列，但是这些转录因子的高亲和性的结合依赖于更长的序列。对于有些转录因子，共有序列有时是不完整的，因为很难确定在4 bp 的核心序列周围哪种碱基组合是高亲和性所必需的。总的来说，TRANSFAC 和 JASPAR 等数据库的质量完全取决于用来建库的实验数据的质量。

利用各种各样的参数进行的数据库搜索，可以使研究人员鉴定出一个可能识别目的序列的 DNA 结合蛋白的集合。这时，研究人员应认真查阅原始文献，来评价定义每个候选蛋白共有序列的相关研究的质量，并评价这些共有序列的完整或不完整的程度。用于定义共有识别序列的结合位点选择研究通常偏向高亲和力的结合位点，从而遗漏了一些以适度减少的亲和力与蛋白质相互作用的 DNA 序列模体，然而这些低亲和力的相互作用在功能上也可以有同样重要的意义。事实上，某些蛋白质的功能性识别序列频繁地以低亲和力相互作用，原因是这些蛋白质需要与结合在邻近位点的蛋白质进行协作结合(见第 1 章和第 11 章)。

该文献综述不应只限于尝试定义共有识别序列的研究，还应涵盖蛋白质-DNA 相互作用的结构研究的仔细评价。由 X 射线晶体学或 NMR（核磁共振）确定的蛋白质-DNA 复合物的结构，可能揭示蛋白质 DNA 结合结构域中特定残基与识别序列的特定碱基之间的关键接触。这些信息有助于确定实验界定的共有序列的有效性及共有序列中特定碱基对的重要性。

鉴定得到一组可能与 DNA 模体相互作用的蛋白质之后，其他一些考虑可以帮助我们导向最吸引人的候选者。一个标准是发现在进化上保守的精确的 DNA 序列或转染实验中控制区的功能活性所必需的精确序列。如果发现对蛋白质-DNA 相互作用重要的核苷酸序列紧密对应于进化保守序列或功能活性所需要的序列，那么候选蛋白与功能元件相关的信心将增加。相反，如果蛋白质结合需要的关键碱基对并不是进化上保守的或与功能性重要碱基对不一致，那么研究人员会预计候选蛋白可能并不是功能相关的。

另一个重要标准是候选蛋白是否在适当的细胞类型中、在适当的发育阶段和（或）在适当的刺激应答中表达。如果该蛋白质在靶基因表达的细胞中不表达，该信息可能仍然是有价值的，因为识别相似 DNA 序列的同一蛋白质家族中的另一个成员可能代替负责靶基因的表达。在这一点上，各种生物的全基因组序列，以及获得表达谱信息的容易性，会非常有价值。如果数据库分析预测到一种特定的蛋白质与 DNA 模体相互作用，那么现在很容易鉴定含有可能与相似的 DNA 序列相互作用的同源 DNA 结合结构域的蛋白质家族的所有成员。此外，可以很容易地进行微阵列实验以鉴定在目的细胞中表达的家族成员。虽然了解一个蛋白质家族的多个成员识别相似的 DNA 序列会有相当大的价值，但是这个事实让人很难精确确定哪个家族成员或哪些家族成员负责该元件的功

能。这个论题在第8章中进行更详细的论述。

最后，除了评价文献、进化保守的碱基对、元件功能需要的碱基对及表达数据，还可以进行EMSA和ChIP实验，以确定预测的蛋白质是否的确能在体外与DNA模体结合，并在体内与模体附近的DNA相结合（见下文和第8章）。

DNA结合蛋白的MudPIT分析

如果在数据库搜索中在预测互作的那些蛋白质之外适于搜索出能识别目的模体的DNA结合蛋白，那么在本章后面描述的大部分常规策略已经在很多研究中取得成功。然而，就如本章的引言中所描述的，基于先进的质谱技术的新策略已经开始成为可行的替代策略，并可能最终取代传统的策略。

历史上，一直难于分离能与目的DNA模体结合的真核蛋白和克隆编码这些蛋白质的基因。最常见的策略一直是通过序列特异性DNA亲和层析纯化蛋白质。在尝试纯化之前，通常重要的是发展蛋白质-DNA相互作用的分析方法以检测该蛋白质。对于亲和层析而言，通常将包含目的DNA模体的DNA片段或寡核苷酸连接成阵列，并共价交联到树脂上。之后，将细胞核提取物或来自较早的传统常规层析柱的馏分加到亲和层析柱上，接着进行洗涤和洗脱。起初，用Edman降解获得通过十二烷基硫酸钠-聚丙烯酰胺凝胶电泳（SDS-PAGE）观察到的纯蛋白的部分氨基酸序列，这导致通过使用简并DNA探针的文库筛选克隆cDNA。几年前，质谱技术在很大程度上取代了Edman降解用于鉴定纯化的DNA结合蛋白。

在大多数情况下，这种传统的蛋白质纯化策略已被证明是极具挑战性的。典型的转录因子是中等丰度的（在每个细胞中25万个拷贝），但通常它们被提取在核提取物中的效率很低。因此，往往难以纯化得到SDS-PAGE凝胶检测所需的足量的蛋白质。此外，细胞中含有大量的DNA和RNA结合蛋白，这些蛋白质很容易与DNA亲和层析柱结合，但没有特异性。这些非特异性核酸结合蛋白使研究人员难以鉴定以序列特异性方式与亲和层析柱结合的转录因子。为了帮助克服这一困难，研究人员一般在上柱之前向提取物中添加非特异性竞争DNA。DNA亲和层析前的传统层析步骤也可以帮助从非特异性结合蛋白中分离目的蛋白。在亲和层析步骤中包含大量竞争性DNA会减少目的蛋白的产量，因为大部分所需的蛋白质与DNA竞争物结合而不是与亲和层析柱结合，通过增加起始材料的量可以克服减产。由于这些挑战（以及下面描述的挑战），纯化DNA结合蛋白可能需要一年或一年多的努力，以及大量蛋白质纯化的专门技能。

重要的是，质谱技术的巨大进步，加上促进肽与基因组编码蛋白配对的完整基因组序列的获得，已导致以更迅速地鉴定与目的DNA模体结合的蛋白质为目标的努力。事实上，由于这些进步，可能很快消除两种需求，一是发展一种初始蛋白质-DNA相互作用分析方法的需求，二是在使用质谱鉴定可与DNA模体结合的蛋白质之前，纯化蛋白质的需求。利用合成的含有目的DNA模体的生物素化双链寡核苷酸（DNA模体两边各有5～10 bp的侧翼序列），就是一个这样的策略。在该策略中，将生物素化的寡核苷酸与适当的核提取物孵育，然后利用链亲和素包被的树脂或磁珠进行一个简单的捕获步骤。经过洗涤后，就可以利用MudPIT来鉴定在沉淀中存在的蛋白质（Washburn et al. 2001；Fourn ier et al. 2007）。之后，研究人员可以通过对收集的蛋白质的排序，确

定哪些蛋白质可能负责 DNA 元件的功能。作为对照，可以在已知对该元件的功能关键的碱基对中含有突变的生物素化的寡核苷酸进行平行实验。结合突变寡核苷酸的蛋白质可能会被排除在候选结合蛋白名单之外。Reed 等（2008）提供了一个成功运用该技术的例子。

这个策略为研究人员提供了几个可能与目的 DNA 模体功能性相互作用的候选蛋白。获得多个候选蛋白既有优势也有劣势。其中一个重要的优势是，得到的结果并不限于能产生可检测的 EMSA 复合物的丰度较高的蛋白质（这些蛋白质不一定是功能相关的）。MudPIT 结果同时揭示能与 DNA 元件结合的丰度较高的和丰度较低的蛋白质。然而，对候选蛋白质进行排序可能需要花费相当长的时间和相当大的精力。基于它们已知的功能或与核酸几乎无特异性的结合的特性，有些蛋白质被鉴定为重要转录因子的可能性非常低。因此，可将注意力集中到那些具有特定结构域（如锌指结构、同源域、Ets 结构域）的蛋白质上，这些结构域通常发现于序列特异性蛋白质-DNA 的相互作用中。之后，可用 ChIP 实验、表达研究和对每个候选蛋白识别的特定碱基对的评价，确定哪些候选蛋白是最切实可行的（第 8 章）。在某些情况下，研究者使用这一 MudPIT 策略可能很快鉴定功能性重要的蛋白质。然而，在其他情况下，对很多候选蛋白的排序需要花费的时间可能比得上，甚至超过纯化 DNA 结合蛋白所花费的时间，这些 DNA 结合蛋白是最初通过 EMSA 鉴定和表征的富有吸引力的 DNA 结合蛋白。

用于粗制细胞裂解物的蛋白质-DNA 相互作用分析方法的发展

虽然正在发展的新的质谱策略不需要起初使用一种体外蛋白质-DNA 相互作用分析方法，但是最终还是需要这样的一种分析方法，以表征被鉴定的候选 DNA 结合蛋白。另外，由于对 MudPIT 分析中鉴定的蛋白质进行排序的挑战性，更可取的可能是用传统 DNA-蛋白质相互作用分析方法追寻 DNA 结合蛋白，之后纯化被鉴定的富有吸引力的 DNA 结合蛋白。最后，即使在最初的 TRANSFAC 和 JASPAR 数据库搜索中提示一个特定的蛋白质可结合该 DNA 模体，蛋白质-DNA 相互作用分析仍然是需要的，以确证和进一步表征预测的相互作用。因此，我们将详细描述这些分析方法的发展。

检测蛋白质-DNA 相互作用的标准方法

已经有几种方法用于体外检测和表征蛋白质-DNA 相互作用，包括 EMSA、DNase I 足迹、外切酶核酸 III 足迹、South-western 印迹、各种化学保护和干扰分析，以及紫外交联。然而，EMSA 是至今最常用的分析方法，这主要是因为它在技术上最简单，并且通常最灵敏。尽管 DNase I 足迹的使用在快速减少，但它仍然是第二种最常用的分析方法，该方法的主要优点是让研究人员能够在较大的 DNA 片段中定位蛋白质结合位点，但是这一优势变得越来越不显著。我们将本章的讨论主要限定在这两种方法上，第 11 章将描述可用来研究蛋白质-DNA 相互作用的其他技术。

*基本的 EMSA 和 DNase I 足迹策略。*在第 11 章中从对表征纯化的重组蛋白感兴趣的研究人员的角度对 EMSA 和 DNase I 足迹进行详细描述。简单地说，EMSA 基于如下原理：蛋白质-DNA 复合物在自然凝胶中比游离 DNA 迁移得更慢，复合物迁移率取决于蛋白质的大小、形状、电荷及多聚化状态（见第 11 章，方案 11.5）。因此，细胞

粗提物中特异性识别给定控制元件的蛋白质可以通过如下方法进行鉴定。首先，将放射性标记的 DNA 小片段（即探针）与提取物进行孵育，使蛋白质-DNA 复合物形成。然后，将混合物通过聚丙烯酰胺凝胶电泳，将游离的放射性标记探针 DNA 和与蛋白质结合的探针分子分离。最后，通过放射自显影或磷屏成像分析对游离型和结合型 DNA 分子进行检测。

对于 DNase I 足迹分析，首先将放射性标记的 DNA 探针和细胞提取物进行孵育，形成蛋白质-DNA 复合物（见专题 11.3，方案 11.1)，这种结合反应和用于 EMSA 的结合反应之间的一项重要区别是：用于 DNase I 足迹分析的放射性标记双链 DNA 探针必须只标记在分子的一端和一条链上（相反，对于 EMSA，放射性标记核苷酸的位置和数目并不重要）。然后，将含有探针和提取物的溶液用极限浓度的 DNase I 处理短的确定的时间。选择 DNase I 处理参数，以便每个探针分子平均只被切割一次。因为这种切割在很大程度上是随机的，因此这种切割将产生不同长度的一套嵌套的 DNA 分子，这些 DNA 分子的长度从单核苷酸和二核苷酸到未被切割的全长分子不等。得到的 DNA 片段先进行变性 PAGE 分析，接着进行放射性自显影或磷屏成像分析。如果所用的 DNase I 消化参数合适，不含细胞提取物的样品会形成从图像底端到顶端的梯状条带。然而，如果提取物中某种蛋白质可以与探针上发现的特定的 DNA 序列相结合，该蛋白质就会保护那个区域的核苷酸不被消化。在结果图片上，相对于没有提取物的反应

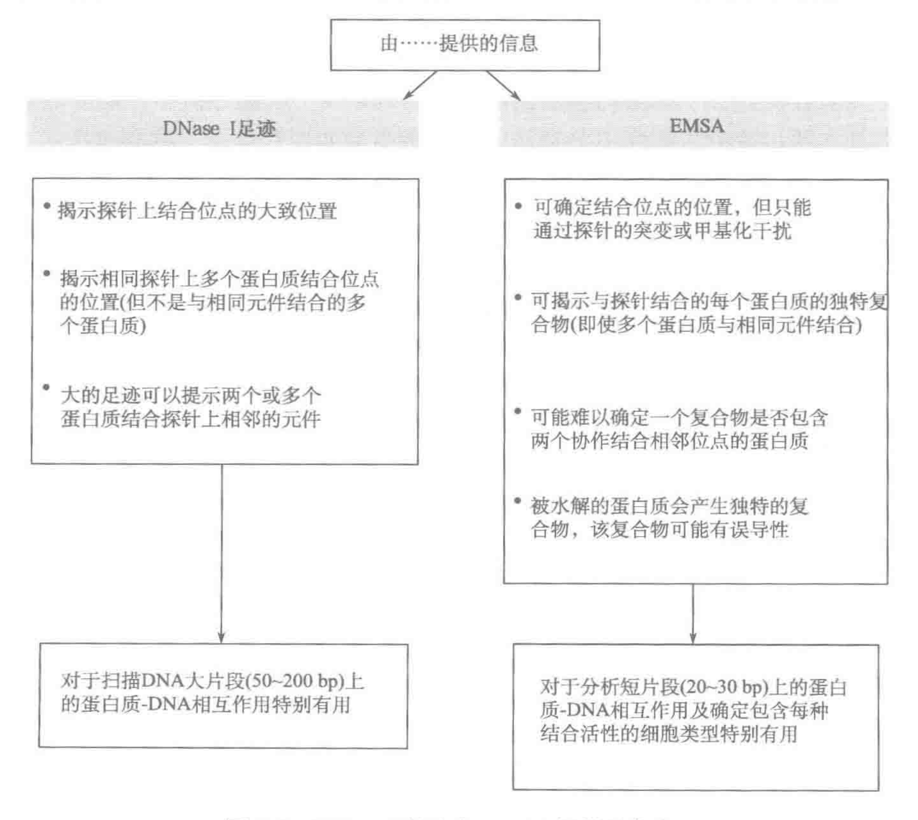


图 7.2　DNase I 足迹和 EMSA 提供的信息。

所产生的条带，这种保护作用会表现为无条带的区域（或包含很弱条带的区域），该受保护的区域称为 DNase I 足迹。在足迹反应物相邻的泳道设置 DNA 测序分子质量标准物，以确定受蛋白质保护的核苷酸。

*通过 EMSA 和 DNase I 足迹的蛋白质差异检测。*虽然许多蛋白质均可以利用 EMSA 和 DNase I 足迹进行检测，但有些 DNA 结合蛋白只能利用这些分析方法中的一种进行检测。对于大多数蛋白质，只能由两种分析方法中的一种检测的例子，尚未确定确切的原因。然而，通过对两种分析方法间主要不同点的概括，可以洞悉可能的解释（图 7.2 和表 7.1）。

表 7.1　EMSA 和 DNase I 足迹分析间的特征型差异

EMSA	DNase I 足迹
当蛋白质处于低浓度时，分析有效	当蛋白质主要处于高浓度时，分析有效
蛋白质-DNA 互相作用必须在凝胶电泳中存活	结合状态必须只在溶液中稳定
大量非特异性竞争物会破坏特异性相互作用	DNase I 活性所需的 $MgCl_2$ 和 $CaCl_2$ 会破坏特异性相互作用

首先，EMSA 灵敏度更高，且即使当提取物中蛋白质浓度较低时，也能揭示特异性的蛋白质-DNA 复合物。其原因是：结合有很少探针分子的蛋白质依然产生比游离型的探针迁移得更慢的复合物。如果探针的比活性足够高，那么通过放射性自显影或磷成像分析就可以很容易地检测该复合物。相反，为了通过 DNase I 足迹检测蛋白质-DNA 复合物，大多数探针分子都必须被蛋白质结合。DNase I 足迹分析中，需要高效率的结合，因为基本的足迹分析程序不分离结合型和游离型的探针分子。在凝胶图像上，条带须显著减少以产生令人信服的足迹。例如，如果仅有 50%的探针分子与蛋白质结合，受保护的条带与不含提取物的条带相比仅减少两倍。对这种小的差异通常很难进行评价。由此可知，由于灵敏度的不同，某些蛋白质，主要是在提取物中以低浓度存在的蛋白质，可能通过 EMSA 是可检测的，但通过 DNase I 足迹则不可检测。

两种分析方法间的第二个主要差异是 DNase I 足迹分析完全在溶液中进行，而 EMSA 则需要在凝胶电泳过程中维持蛋白质-DNA 相互作用。几乎没有蛋白质以与一般 2～4 h 的电泳时段相兼容的解离速率与 DNA 结合。对于大多数蛋白质来说，EMSA 的成功依赖于一种被称为“笼合”的过程，在该过程中，当蛋白质和 DNA 在电泳期间解离时，凝胶介质会维持蛋白质与 DNA 相互紧密靠近，从而使它们快速重新结合，并作为复合物迁移通过凝胶。虽然笼合过程有助于对许多蛋白质-DNA 复合物的分析，但有些复合物似乎非常难维持。复合物不能耐受凝胶电泳的证据通常可通过如下观察来提供，即在 DNase I 足迹分析中能观察到该蛋白质产生强保护作用，而相同量的蛋白质和探针只显示微弱的或不可检测的 EMSA 复合物。

EMSA 和 DNase I 足迹分析之间的第三个差异是值得注意的，即用 EMSA 分析细胞粗提物时非特异性竞争 DNA 的用量。在两种分析中，通常加入过量的竞争 DNA，如 poly（dI：dC）或 poly（dA：dT）等，以防止非特异性和低特异性 DNA 结合蛋白与探针结合并掩盖特异性复合物。对于 DNase I 足迹反应来说，10 ～100 μg 细胞粗提

物中通常加入 1 μg poly（dI∶dC）。相比之下，对于含有 5 ～10 μg 粗提物的 EMSA 反应来说，通常需加入 2 ～8 μg poly（dI∶dC）。EMSA 反应需要相对高浓度的竞争性 DNA，因为适量的非特异性核酸结合蛋白会导致凝胶图像上产生放射性探针污迹，这种污迹阻碍目的复合物的检测。对于 DNase I 足迹分析来说，非特异性结合蛋白分布在整个探针上，并且因此不太可能显著影响特异性蛋白质-DNA 复合物的检测。如果目的 DNA 结合蛋白对竞争 DNA 具有中等亲和力，这种差异就会影响每次分析成功的程度；用于 EMSA 反应中的高浓度竞争性 DNA 可能会与探针竞争性地与目的蛋白质相结合，从而减弱或消除特异性蛋白质-DNA 复合物。第四个差异是 DNase I 足迹分析中为了 DNase I 的活性需要加入 $MgCl_2$ 和 $CaCl_2$。虽然这些化学物质产生的离子不影响大多数蛋白质-DNA 的相互作用，或者促进蛋白质-DNA 的相互作用，但也有可能削弱某些相互作用。

根据这些差异，就不会奇怪某些 DNA 结合蛋白的活性更容易以两种方法中的一种方法所鉴定。若一种蛋白质仅被 EMSA 检测，通常是该蛋白质浓度不足以与 DNase I 足迹分析中所需的高比例的探针分子相结合。例如，通过 EMSA 比通过 DNase I 足迹更容易检测核因子-κB（NF-κB），因为 NF-κB 在核提取物中浓度相对较低。相比之下，Spl 蛋白和 Ikaros 蛋白最初比较容易被 DNase I 足迹分析所检测，直到发展了相关条件，它们才成功地被 EMSA 检测（Dynan and Tjian 1983；Lo et al. 1991；Hahm et al. 1994）。

对于大多数研究人员来说，测试各种各样 EMSA 分析的条件比利用 DNase I 足迹作为日常分析要更好。DNase I 足迹早期的一个优势是在搜索可结合大的启动子或增强子片段的蛋白质时，它提供了该相互作用在探针上的位置的信息。而现在，研究人员一般使用短的合成的寡核苷酸，这就不再是个大问题了。但当研究人员发现需要搜寻可与相对长的 DNA 片段结合的蛋白质，或者研究人员有兴趣确定一个纯蛋白是否能协作性地与一个太大而无法用于 EMSA 实验的 DNA 片段上的多重位点结合时，DNase I 足迹分析仍然是很有用的（见第 11 章）。

电泳迁移率变动分析法

如上所述，EMSA 是用于研究蛋白质-DNA 相互作用的常用且相对简单的方法。在第 11 章（方案 11.5）中将从分析纯重组蛋白的角度对此方法作详细的描述。这里，我们重点讨论当使用 EMSA 初步鉴定和表征粗提物中与某一明确的 DNA 元件结合的蛋白质时，必须考虑的具体问题。

*放射性标记探针的设计。*放射性标记 EMSA 探针的设计会影响所得到的结果的质量。对于在粗提物中检测到的许多蛋白质来说，约 20～25 bp 的双链寡核苷酸可以产生一致的、高质量的结果。例如，通常使用 19 bp 寡核苷酸进行检测 NF-κB 复合物的 EMSA 实验（Pierce et al. 1988）。寡核苷酸探针在已确定的功能控制元件的两端应至少含有 5～10 bp 的序列，因为这些邻近的碱基虽然不被蛋白质特异性识别，但对于稳定的蛋白质-DNA 相互作用可能是必需的。

为了准备双链片段，可合成并退火两条单链的寡核苷酸（Ausubel et al. 1994，单元 6.4），或者，可准备含有识别序列的两个互补序列的长链寡核苷酸，使该寡核苷酸

自我退火，形成柄环结构。可对粗制脱盐的合成寡核苷酸进行放射性标记，具体方法是以［γ-^{32}P］ATP 和 T4 多核苷酸激酶对其 5′端进行放射性标记（方案 3.1，Sambrook and Russell 2001，p. 7.78），或使用 α-^{32}P 标记的核苷三磷酸和大肠杆菌（*Escherichia coli*）DNA 聚合酶 Klenow 片段自 3′缩进端进行聚合（Ausubel et al. 1994，单元 3.5；Sambrook and Russell 2001，p. 10.30）。只有当核苷酸已退火形成双链分子时，使用后一种方法进行标记才有优势，因为 Klenow 片段掺入核苷酸需要模板。

对于某些 DNA 结合蛋白来说，20～25 bp 的短核苷酸并不能支持稳定的蛋白质-DNA 相互作用，因此有必要使用更长的探针。短探针的失败可能产生于蛋白质为了稳定结合需要较长的一段邻近 DNA。或者，含有特定识别序列的短探针可能未正确退火、在结合反应期间解链或具有阻碍蛋白结合的二级结构。可以使用几种方法生产较长的探针，其中一种方法是简单地使用较长的合成寡核苷酸。另一种方法是从含有 DNA 元件的质粒中切出并放射性标记一个限制性片段（方案 11.6）。

也可以用使用放射性标记引物的聚合酶链反应（PCR）制备 EMSA 探针（Ausubel et al. 1994，单元 15.2）。PCR 的一个优点是，如果在野生型和突变型两个模板上使用相同的放射性引物，则可以制备具有特异活性的野生型和突变型探针。如果为一个转染分析中的突变体分析已经生产了一个控制元件中的一系列突变体，则可以使用相同的两个 PCR 引物为 EMSA 实验制备含有野生型序列和每个突变型序列的探针。

提取物的制备。绝大多数体外蛋白质-DNA 作用的研究都使用按照 Dignam 等（1983）的方法制备的核提取物。该方法在第 12 章（方案 12.1）中被详细描述，简单地说，就是将细胞在低渗缓冲液中膨胀，然后用 Dounce 匀浆器通过物理破碎使细胞裂解。另一个可选的方法是使用低浓度的非离子型去污剂 NP-40 使细胞裂解。不论是物理裂解，还是去污剂裂解，来自大多数哺乳动物细胞类型的细胞核保持相当的完整，并且可以通过低速离心从细胞质蛋白中分离出来。为了从细胞核中提取大多数 DNA 结合蛋白，可将细胞核沉淀重悬和搅拌在含高浓度 NaCl 或 KCl 的缓冲液中。Dignam 和 Roeder 根据经验确定，0.42 mol/L 的 NaCl 浓度对于制备核提取物最佳，所制备的核提取物支持从腺病毒主要晚期启动子的体外转录。大多数研究人员也利用这一盐浓度制备用于 DNA 结合蛋白分析的核提取物。然而，当追寻新的 DNA 结合活性时，改变用于提取物制备的盐浓度是有益的，较低的盐浓度能防止酶抑制剂和非特异性结合蛋白的提取，从而增加 EMSA 分析的成功率。目标蛋白的有效提取可能需要较高的盐浓度。

如果通过 EMSA 检测 DNA 结合活性的尝试失败了，可以进一步浓缩或分级核提取物，以增加成功的可能性。可以通过硫酸铵沉淀及其随后的彻底透析增加提取物的蛋白质浓度（见第 12 章，方案 12.1）。使用更加浓缩的提取物可能会更容易检测目的 DNA 结合活性。Dignam 和 Roeder 发现使用 53%的硫酸铵沉淀的提取物对提高体外转录提取物的活性是最佳的，但是对于特定的 DNA 结合蛋白来说，硫酸铵的最佳百分比可能有所不同。

改善 DNA 结合活性检测的第二种策略是将提取物分级（见第 12 章）（Ausubel et al. 1994，单元 10.10）。带负电荷的亲和树脂、肝素琼脂糖和肝素葡聚糖对于这个目的是有用的。如果将提取物加到含有这些在 0.1mol/L KCl 或 NaCl 的缓冲液中的树脂的

柱子上，使用0.4 mol/L的盐梯度将洗脱大多数DNA结合蛋白。当蛋白质以异常高的亲和力和树脂结合时，也应该进行1 mol/L的盐梯度洗脱。分级程序将DNA结合蛋白从其他可能抑制结合反应的蛋白质中分离出来。此外，特定蛋白质在0.4 mol/L和1 mol/L盐分步收集物中的浓度可能比粗提物中的浓度更高。因此，在这些分步收集物中可能检测到那些在粗提物中检测不到的DNA结合活性。

如果使用核提取物检测候选结合活性的所有尝试都失败了，则可以测验全细胞提取物、细胞质提取物或其他核提取物制备方法（见第12章）（Gorski et al. 1986；Shapiro et al. 1988；Soeller et al. 1988）。在核提取物制备期间，某些细胞核蛋白泄漏进细胞质分步收集物中，从而导致在细胞质提取物或全细胞提取物中DNA结合活性可能更丰富。

*结合反应中的竞争性DNA和其他组分。*加入结合反应中的非特异性竞争DNA对于EMSA实验的成功会是至关重要的，特别是当使用细胞核粗提物时。竞争DNA阻止非特异性和低特异性核酸结合蛋白与放射性标记的探针稳定地相互作用。在没有合适的非特异性竞争性DNA的情况下，提取物中的许多蛋白质（包括高丰度的RNA结合蛋白）都会与探针结合，导致凝胶图像上出现放射性污迹。增加竞争性DNA的浓度，会使更多的蛋白质分子与放射性标记的探针相隔离，结果是主要留下那些相对于非标记的竞争性DNA对探针有最高特异性亲和力的蛋白质。虽然高浓度竞争性DNA通常会提高结果的质量和清晰度，但它们也会增加相关性DNA结合蛋白和标记探针相隔离的可能性。

使用最频繁的非特异性竞争性DNA是poly（dI∶dC），然而，其他一些竞争性DNA，包括poly（dG∶dC）、poly（dA∶dT）、经剪切的*E. coli*或鲑鱼精基因组DNA，可能产生更好（或者至少不同）的结果。如果目的结合蛋白以高亲和力与竞争性DNA结合，在EMSA凝胶上就可能不会产生可检测的复合物。相反，如果那些蛋白质与竞争性DNA以低亲和力结合，可能会观察到与控制元件的功能不相关的包含蛋白质的复合物。因此，在尝试鉴定DNA元件的相关调控因子的候选蛋白时，测验不同的竞争性DNA是有益的。

除了竞争性DNA，结合反应体系中的其他组分也可作变动，以加强蛋白质-DNA复合物的检测，如单价或二价阳离子的浓度。结合反应中加入低浓度（如0.01%）的非离子型去污剂NP-40可能对某些蛋白质-DNA复合物有益。而聚乙烯醇（2%）的存在可能对其他一些复合物有益，聚乙烯醇能增加提取物中蛋白质的有效浓度。优化EMSA实验结合反应的其他建议参见方案11.5和文献Ausubel等（1994，单元12.2）。

*凝胶电泳实验的条件。*当对目的DNA元件发展一种EMSA分析时，强烈建议以不同的凝胶电泳缓冲液做试验性实验，因为有些蛋白质-DNA复合物的检测依赖于所用的特定缓冲液。最常用的三种EMSA凝胶缓冲液是Tris-硼酸-EDTA（TBE）、Tris-乙酸-EDTA（TAE）和Tris-甘氨酸。甘油（5%）的存在也会提高某些结合活性的检测。另外，可以在室温或4℃下进行电泳。有关不同凝胶电泳缓冲液的更多信息参见第11章中的方案11.5和文献Ausubel等（1994，单元12.2）。

*初始实验的设计和结果的解释。*以包含目的控制元件的放射性标记探针进行的初始

EMSA 实验，可能会产生检测到一个或多个蛋白质-DNA 复合物的结果。初始实验通常以几种不同量的提取物进行（2～20 μg），而且，对于每一种量的提取物应使用不同量的非特异性竞争性 DNA（1～10 μg）。如果未检测到复合物，可以根据上面提到的方法改变结合反应和凝胶条件。对于所检测到的每种复合物，最初需要解决两个问题：① 蛋白质-DNA 复合物对所测验的探针特异吗？② 该复合物有很大的可能性负责在内源控制区环境中该 DNA 元件的体内功能吗？

这些问题一般通过在特异性竞争 DNA 存在的情况下或以突变体探针进行另外的 EMSA 实验来解决。通过比较使用放射性标记野生型探针和突变型探针所获得的复合物，通常会获得最直接的信息。如果在目的控制元件中有核苷酸替换的探针未观察到蛋白质-DNA 复合物，那么该蛋白质-DNA 相互作用几乎肯定需要那些被替换的核苷酸。通过表明该相互作用是序列特异性的并需要功能性相关的核苷酸，这一结果就回答了或至少部分地回答了上面提出的两个问题。通过制备含有其他替换突变的探针，可获得该复合物可能负责控制元件功能的更多的证据。如果这种选择性扰乱调控元件功能（如转染实验中的功能）的核苷酸替换与扰乱 EMSA 复合物的核苷酸替换相一致，则该复合物是功能相关的可能性将增大（第 8 章）。

正如所有的标准 EMSA 实验，应以低浓度的探针（1～10 fmol/反应）和能在线性范围内产生复合物的蛋白质浓度（即 1%～50%的探针和蛋白质相结合）进行突变分析。使用野生型和突变型探针的比较应以具有相似特异活性的放射性标记探针进行。如果突变型探针比野生型探针的比活性低，即使突变对相互作用的亲和力没有影响，蛋白质-DNA 复合物的丰度也不高。不幸的是，蛋白质对野生型和突变型序列的相对亲和力不能简单地通过比较使用野生型探针和突变型探针所形成复合物的丰度而定；如果该复合物不是足够稳定以至于在凝胶电泳中不能生存，亲和力的轻微降低也可导致复合物量的急剧减少。因此，如果一个突变使复合物的丰度减小，只能推断该蛋白质对突变体探针的亲和力比对野生型的低，差别的大小无法辨别。此外，如果一个突变对复合物的形成无影响，该蛋白质很可能是以相当的亲和力和野生型及突变型探针结合。总的来说，尽管缺少定量的数据，但野生型和突变型探针的比较也可提供以下有价值的信息：① 一个相互作用的特异性；② 该相互作用所需要的核苷酸；③ 该相互作用所需核苷酸与 DNA 元件功能所需核苷酸间的关系。

检查蛋白质-DNA 复合物的特异性和功能相关性的另一种方法是以未标记的野生型和突变型 DNA 片段，如退火的寡核苷酸、PCR 产物或限制性片段，进行竞争性实验（见专题 7.1）。为了使竞争性实验的结果有意义，必须对野生型和突变型竞争性 DNA 进行仔细定量，以确保它们浓度相当。通过在未标记野生型或突变型竞争性 DNA 浓度逐渐升高的一系列结合反应中加入放射性标记的野生型探针，来进行竞争实验。在未标记竞争性 DNA 存在的情况下，放射性标记复合物的丰度减弱，因为 DNA 结合蛋白使自己随机地分布于探针和竞争性分子中。由于野生型竞争性 DNA 对蛋白质的亲和力比突变型竞争性 DNA 更高，因此降低该复合物的丰度只需较低浓度的野生型竞争性 DNA。

通过将野生型竞争性 DNA 与一系列突变型竞争性 DNA 的比较（专题 7.1），可以

获得有关蛋白质-DNA 相互作用的序列特异性和功能相关性的宝贵信息。如果野生型竞争性 DNA 对 EMSA 复合物的减弱程度比突变型竞争性 DNA 的大，则该蛋白质一定以序列特异性的方式与 DNA 结合。如果减弱竞争的突变体对应于减弱 DNA 元件功能（如在转染分析中）的突变体，则提示该蛋白质-DNA 相互作用可能是功能相关的。

专题 7.1　假设 EMSA 结果

在正文中描述了 EMSA 竞争实验的基本概念。图 7.3 中的假设竞争实验说明了这些概念。在没有特异性竞争性 DNA 存在的条件下，利用提取物和野生型探针（泳道 1），获得了 4 个复合物（A～D）。该结合反应中包含了大量竞争性 DNA，如 poly（dI：dC），以减少非特异性复合物的形成。

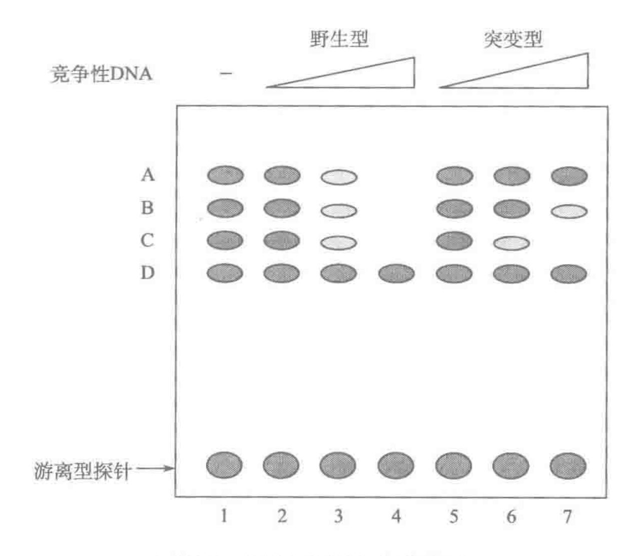


图 7.3　假设 EMSA 竞争实验。

当野生型竞争性 DNA 的浓度增高时，复合物 A 消失（泳道 2～4）。相比之下，复合物 A 的丰度不受突变型竞争性 DNA 的影响（泳道 5～7）。该结果提示负责复合物 A 的蛋白质是一种序列特异性 DNA 结合蛋白，它对突变型竞争性 DNA 的亲和力远低于对野生型竞争性 DNA 的亲和力。如果突变体中所改变的核苷酸已知对控制元件的功能重要（如在转染分析中），则上述结果提示该蛋白质至少与探针中具有重要功能的部分核苷酸相结合。

野生型竞争性 DNA（泳道 2～4）和突变型竞争性 DNA（泳道 5～7）都减少复合物 B，但将复合物的丰度减小到特定的量需要较高浓度的突变型竞争性 DNA（比较泳道 2 和泳道 7），这个结果比用复合物 A 观察到的结果更有代表性。像大多数序列特异性 DNA 结合蛋白一样，负责复合物 B 的蛋白质与一系列广泛的 DNA 序列以低亲和力结合。因此，要减少特异性蛋白质-DNA 复合物，需要相对高浓度的突变型竞争性

DNA是很正常的。尽管如此，复合物B的行为提示，像复合物A一样，它含有与功能重要核苷酸接触的序列特异性DNA结合蛋白。

野生型竞争性DNA和突变型竞争性DNA对复合物C的减少程度相似。该结果提示蛋白质的结合不需要突变型探针中改变了的核苷酸。尽管如此，该蛋白质仍可能是识别两个探针分子上另一段序列的一种序列特异性DNA结合蛋白。对该结果的另一解释是此蛋白质非特异性地结合DNA，但这不太可能。因为非特异性DNA结合蛋白预计会与任何放射性标记探针结合。通过检验具有无关DNA序列的大量标记探针，可以检查这种可能性。另外，也可以检验目的探针的另外的突变体。

像以复合物D为研究对象获得的这一类结果通常被误解；野生型或突变型竞争性DNA都不能减少这种复合物。许多已发表的论文都建议这种行为表明该复合物含有非特异性DNA结合蛋白。相反，该结果仅表明该蛋白质在提取物中丰度异常高，并且可能与探针以相对低的亲和力结合，未提供任何该蛋白质序列特异性的有意义的信息。不论该蛋白质是否是序列特异性DNA结合蛋白或非特异性核酸结合蛋白，通过加入足量的非标记野生型竞争性DNA减少该复合物都是可能的。如果该复合物没有减少，该结果仅仅表明需要更高浓度的特异性竞争性DNA（或更低浓度的提取物）。也可以通过分析该蛋白质结合含有突变或无关序列的大量放射性标记探针的能力来阐明其序列特异性。怀疑复合物D为非特异性核酸结合蛋白的原因是许多这样的蛋白质含量丰富，而大多数典型的转录激活因子和抑制因子在细胞和细胞提取物中以相对低的浓度存在。因此，当相当大的过量野生型竞争性DNA未减少一个复合物时，该复合物有可能不是转录激活因子或抑制因子。尽管如此，当研究人员开始从事对染色质结构和核酸组织的研究时，他们对高丰度、序列特异性DNA结合蛋白的兴趣可能会增大。

利用突变型竞争性DNA而不用放射性标记突变型探针评价特异性的主要优点是，前者提供关于蛋白质对野生型和突变型序列相对亲和力的更准确的信息。如上所述，放射性标记野生型和突变型探针的比较提供了关于亲和力的有限的信息。而对竞争实验则不需要给出相同的警告，因为在最初的结合反应中，涉及野生型和突变型序列比较的关键事件发生在溶液中，凝胶电泳则是在竞争完成之后才进行的。如果蛋白质对突变型序列的亲和力比对野生型序列的亲和力小两倍，那么两倍于野生型竞争性DNA浓度的突变型竞争性DNA应将复合物的丰度减小到特定的量。因此，竞争策略更有可能提供关于相对亲和力的有意义的定量信息。

虽然上述考虑看起来很简单，但竞争实验的结果可能是复杂的。例如，复合物丰度的两倍减少通常需要大量过剩的野生型竞争性DNA。这是因为尽管在凝胶的底部存在游离型探针，但大多数EMSA实验都含有过量的蛋白质。因此，只有当竞争性DNA的浓度足以超过相互作用的K_d时，竞争才会发生。在第11章中对K_d的测量有进一步的讨论。虽然这个问题使得对特定浓度的野生型竞争性DNA的效应难以预测，但其并未使野生型和突变型竞争性DNA的比较完全无效。即使需要50倍过量的野生型竞争性DNA才得到蛋白质-DNA复合物丰度的两倍减小，通过突变型竞争性DNA的效应可洞悉蛋白质对野生型和突变型序列的相对亲和力。若50倍过量的突变型竞争性DNA

引起蛋白质-DNA 复合物丰度相似的两倍减少，则可以断定该突变对蛋白质和 DNA 的亲和力没有显著影响。如果为了使蛋白质-DNA 复合物的丰度减小两倍，需要 100 倍过量的突变型竞争性 DNA，则该蛋白质对突变序列的亲和力比对野生型序列的亲和力大约小两倍。由于用于评价蛋白质-DNA 相互作用特异性和功能相关性的放射性标记探针和竞争策略具有独特的优势及局限性，因此强烈建议同时使用两者。

通过甲基化干扰分析 EMSA 复合物。上面讨论的实验提供了通过 EMSA 检测蛋白质-DNA 相互作用的特异性及核苷酸需求的相关信息。另一项技术，即甲基化干扰法（在第 11 章和方案 11.4 中详细描述）可以与 EMSA 联用，提供参与相互作用的特定核苷酸接触的信息。简单地说，该分析始于仅在一端标记的 EMSA 探针，该探针标记并纯化后，其鸟嘌呤碱基用硫酸二甲酯修饰。为使该分析成功，蛋白质结合位点一般需要离放射性标记末端至少 25 bp 或 30 bp，结合位点之后还需要 15～20 bp 的序列。因此，用于甲基化干扰的放射性探针一般比用于常规 EMSA 的探针长。选择硫酸二甲酯的浓度和孵育时间，以便平均每个探针分子产生一个甲基化的鸟嘌呤。修饰过的探针和提取物在常规 EMSA 结合反应中孵育。DNA 结合蛋白随机结合修饰的及未修饰的探针分子。但是，如果探针中与蛋白质紧密接触的鸟嘌呤被甲基化，蛋白质可能就不能和探针分子结合。使用 EMSA 凝胶分离游离型探针分子与蛋白质-DNA 复合物，随后放射自显影，将自显影胶片和凝胶对齐，切出含有游离型探针和蛋白质-DNA 复合物的聚丙烯酰胺凝胶薄片。将 DNA 从两凝胶薄片中洗脱出来，之后与哌啶孵育，哌啶在含有甲基化鸟嘌呤的位置上切割每个探针分子。切割产生的 DNA 片段通过聚丙烯酰胺凝胶电泳进行分析，然后进行放射自显影或磷成像分析。如果特定鸟嘌呤的甲基化阻止 DNA 的结合，那么在该鸟嘌呤处切割的 DNA 片段将在来源于蛋白质-DNA 复合物的样品中缺失，但在来源于游离型探针的样品中含量丰富。由于甲基基团十分小，由特定甲基化鸟嘌呤残基引起的蛋白质结合抑制则说明蛋白质和该鸟嘌呤紧密靠近。

在许多方面，甲基化干扰结果可提供与 EMSA 竞争及突变研究所提供的信息相似的信息，这些结果可表明结合是序列特异性的，并能揭示结合位点的位置。甲基化干扰技术的一个局限性是它只提供有关蛋白质-DNA 相互作用所需核苷酸的部分信息。仔细的突变分析能提供更详细的信息，这些信息可与功能分析的结果更有效地相关（见第 8 章）。另外，如果蛋白质-DNA 复合物不丰富，甲基化干扰可能在技术上具有挑战性。对于弱复合物也许不可能获得有说服力的结果，从而使得竞争和突变研究必不可少。对于强复合物，建议同时使用这三种方法，以提供尽可能完整的蛋白质-DNA 相互作用图景。第 11 章将描述可用于蛋白质-DNA 相互作用更严格建模的其他技术。但这种详细的分析通常是在编码该蛋白质的基因已被鉴定后利用重组蛋白进行。

EMSA 复合物中以前已描述过的蛋白质的分析。在已经确定了蛋白质-DNA 相互作用的序列特异性和核苷酸需求之后，通常重要的是要阐明该蛋白质负责之前已描述的相互作用的可能性。这种分析的第一步是搜索结合位点数据库（图 7.1），以确定 DNA 元件是否与一个或多个已知蛋白的识别序列相似。如果通过数据库搜索鉴定了候选蛋白质，可以利用抗体评价它们负责 EMSA 复合物的可能性。许多针对已知 DNA 结合蛋白的抗体可以从多种商业及学术渠道获得。否则，就要利用合成肽或细菌表达的融合蛋

白作抗原制备抗体（Harlow and Lane 1999）。

为了确定一个 EMSA 复合物是否含有已知蛋白质，可以将抗体直接加入到结合反应中，或在加入放射性标记探针之前和提取物预孵育。如果抗体识别 DNA 结合结构域，预孵育会促使其更有效地形成抗体-抗原复合物。在这种情况下，抗体可能会阻止 DNA 结合，结果导致在 EMSA 凝胶图像上观察到蛋白质-DNA 复合物的量减小（图 7.4A）。如果抗体识别不同于 DNA 结合结构域的其他结构域，抗体很可能会“超迁移”复合物，这意味着该复合物将在 EMSA 凝胶上迁移得更慢，因为当与抗体结合时，该复合物的分子质量变得更大（图 7.4B）。通过将蛋白质稳定在一种有能力结合的构象状态，抗体也可以稳定蛋白质-DNA 的相互作用。这导致超迁移的条带，该条带比抗体不存在时观察到的原始条带丰度更高。

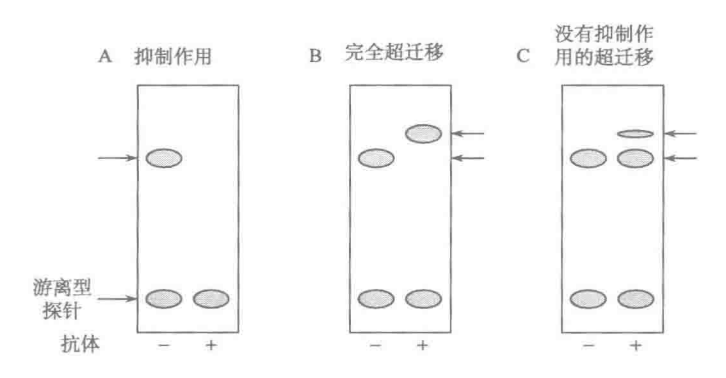


图 7.4 抗体对 EMSA 复合物的影响。

抗体对于确定 EMSA 复合物中蛋白质的身份很有用。但是对其使用要谨慎，主要是因为大多数多克隆及单克隆抗体制剂是不纯的，抗体制剂中的杂质和特异性抗体-抗原相互作用一样能有效地抑制复合物的形成。必须包含仔细设计的对照，特别是当抗体抑制复合物，而不是抑制超迁移时。多克隆抗体的恰当对照是免疫前血清或针对无关蛋白的多克隆制剂。测验一种抗体制剂对一个以无关探针结合该抗体应不识别的蛋白质得到的 EMSA 复合物也是有益的。通过蛋白 A 或蛋白 G 琼脂糖层析或抗原亲和层析纯化抗体，有时可以大大提高抗体效应的特异性（Harlow and Lane 1999）。如果纯化的多克隆抗体和 EMSA 实验联用，对照抗体也应经历同样的纯化方案，并与测验抗体一同分析。如果使用单克隆抗体，应该包含一个对照单克隆抗体，该单克隆抗体是以同样的方法针对一个不相关蛋白制备的。

如果是抗体超迁移复合物，而不是抑制复合物，上面描述的对照就不太重要，但是，也必须同样谨慎地解释超迁移结果。有些情况下，即使凝胶图像上出现弱的或中等强度的超迁移复合物（图 7.4C），蛋白质-DNA 复合物的丰度也不会被该抗体降低。这种结果并未证实该抗体识别的蛋白质是原始复合物的组分。因为一些抗体稳定蛋白质-DNA 相互作用（见上文），超迁移复合物可能对应于一种稳定型的复合物，这种复合物在没有抗体存在的条件下检测不到，且与原始的蛋白质-DNA 复合物不相关。只有当随着超迁移复合物的出现，该复合物的丰度降低时，才能提供一个复合物含有特定蛋白质

的证据。抗体超迁移的另一个需要注意的是，抗体可促进一种蛋白质和两个相邻位点的协同结合。

如果该抗体不能扰乱或超迁移该复合物，被该抗体识别的蛋白质可能不存在于该复合物中。或者，抗体可能太稀或可能以低亲和力与该靶蛋白结合。阳性对照是必需的，在阳性对照中该抗体被表明扰乱或超迁移已知含有该蛋白质的 EMSA 复合物。

如果该蛋白质-DNA 复合物令人信服地被抗已知蛋白的抗体扰乱或超迁移，则该复合物很有可能包含该蛋白质。为进一步验证这种假设，可以获得并分析这种蛋白质的 cDNA。通常通过体外转录/翻译（Ausubel et al. 1994，单元 10.17）生产全长蛋白质，并与细胞粗提物一同进行 EMSA 测验。如果利用粗提物检测到的复合物仅含有一种蛋白质，那么该蛋白质应与含有体外翻译蛋白的 EMSA 复合物共迁移，否则，可能需要重新评价该抗体结果。或者，该蛋白质可能与另一种蛋白质结合形成异二聚体，或者可能与识别探针上邻近位点的另一种蛋白质协同结合。根据对该蛋白质家族一般特征的了解，以及对探针上复合物形成所需核苷酸的详细了解，应使这些可能性得到评价。

如果这些结果提示一个已知蛋白质结合目的控制元件，就可以评价该蛋白质对该元件功能的相关性（见第 8 章）。相比之下，如果结果提示该蛋白质以前没有被描述过，那么可以进行该蛋白质及其编码基因的鉴定。最后一种可能性是该复合物含有一个已知蛋白与一新蛋白的异二聚体。为了鉴定该未知组分，可能需要纯化形成该复合物的蛋白质。

DNase I 足迹

前文简要描述了 DNase I 足迹分析，第 11 章将会进行详尽的阐述。这一部分将讨论利用粗提物进行蛋白质-DNA 相互作用 DNase I 足迹分析时应考虑的具体问题。

*放射性标记探针的设计。*设计探针时应注意两个要点：放射性标记探针的长度及探针上相关 DNA 元件的位置。探针必须在一条链的一个末端进行放射性标记。可以通过直接标记限制性酶切片段（方案 11.6）制备这种探针，也可以利用一个放射性标记的引物和一个未经标记的引物通过 PCR 制备这种探针（Ausubel et al. 1994，单元 15.2.6）。EMSA 常用的人工合成的寡核苷酸探针，不太常用于 DNase I 足迹，因为 DNase I 足迹一般需要较长的探针。

用粗提物进行 DNase I 足迹，其探针长度一般在 100 到几百个碱基对（如第 11 章所述，短的探针通常与重组蛋白一起使用）。重要的是目的 DNA 元件不要在探针两个末端的 40 bp 以内，这种限制的原因之一是一些提取物含有大量的 DNA 末端结合蛋白，这些 DNA 末端结合蛋白可以使在探针末端的特异性蛋白质-DNA 相互作用变得模糊不清。另外，对 DNase I 足迹结果的正确评价依靠对不受蛋白质结合影响的保护区两侧切割产物的检测；如果保护区太靠近探针的末端，在这一末端就难以检测到不受影响的切割产物。

为了准确评价 DNase I 足迹数据，如果目的 DNA 元件大约在标记末端的 200 bp 以内也是最有益的。500 bp 或更长的探针可成功地用于足迹分析，但是发生在距离放射性标记超过 200 bp 的蛋白质-DNA 相互作用通常在足迹分析中难以检测，这归因于凝胶顶部条带的压缩。考虑到这些限制，DNase I 足迹程序通常能提供在大约 160 bp 区域

内发生的蛋白质-DNA 相互作用的有关信息，相当于距离探针标记末端大致 40～200 bp 的区域。因此，如果要分析的调控区长度为 300 bp 左右，仅用两个探针就可能扫描特异性蛋白质-DNA 相互作用的整个区域。

为了进一步增加目的蛋白质-DNA 相互作用将被检测到的可能性，可以测验标记在另一条链上的独立的一套探针。对于某些蛋白质-DNA 相互作用，利用一条链上作标记的探针比利用另一条链作标记的探针更易于检测。另外，同时利用两条链进行分析可减小重要的蛋白质-DNA 相互作用发生在对 DNase Ⅰ 切割有抗性的区域上的可能性。DNase Ⅰ 切割双链 DNA 片段的一条链，这种切割发生在裸 DNA 探针的很多位点上，但并不是完全随机的。因此，可以观察到对 DNase Ⅰ 有抗性、长度为 20 bp 或更长的区域。如果相关的蛋白质-DNA 相互作用发生在这种区域，就无法进行检测。通过测验标记在两条对立链上的两个探针，由于在一条链或另一条链上 DNA 的长的伸出对 DNase Ⅰ 切割具有抗性，遗漏一个相互作用的可能性将会减少。

重要的是要注意 DNase Ⅰ 足迹检测到的蛋白质-DNA 相互作用的位置，是通过比较凝胶图像上受保护区域和在相邻泳道上电泳的 DNA 分子质量标准物进行确定的（见第 11 章）。放射性标记的限制性酶切片段可以提供一些关于探针标记末端到受保护区域的距离的指示，而 DNA 测序分子质量标准物能提供更准确的信息。

提取物准备。准备方法及为 DNase Ⅰ 足迹分析优化和浓缩提取物的策略与 EMSA 实验相似。然而，对于 DNase Ⅰ 足迹而言，更多的时候有必要通过硫酸铵沉淀法或肝素-琼脂糖层析法（见第 12 章）浓缩提取物，这是因为为了检测到足迹，更大比例的探针分子必须被特定蛋白结合（如前所述）。

结合反应。一般来说，利用粗提物进行 DNase Ⅰ 足迹的结合反应与那些用于 EMSA 的结合反应相似，但有些组分的浓度有所不同。在 DNase Ⅰ 足迹分析中，非特异性竞争 DNA，如 poly（dI∶dC）的浓度一般较低，而提取物的浓度一般较高。利用粗提物进行 DNase Ⅰ 足迹时，poly（dI∶dC）的标准浓度是每 50 μl 结合反应 1 μg。最适粗提物的浓度范围是每 50 μl 结合反应 50～250 μg。尤其重要的是需要测验不同浓度的提取物。太高的浓度可导致提取物中蛋白质的探针降解或脱磷酸化或探针对 DNase Ⅰ 切割的完全抗性。

第 11 章中将从分析纯化的重组 DNA 结合蛋白的角度描述 DNase Ⅰ 切割反应。分析粗提物中的 DNA 结合活性时，唯一值得注意的不同点是可能需要更高浓度的 DNase Ⅰ。因此，必须分别确定在缺少蛋白质的对照反应和在含有各种浓度提取物的反应中 DNase Ⅰ 的最适浓度。当反应中提取物的浓度增加时，可能需要进一步增加 DNase Ⅰ 浓度。对于每一组反应条件而言，目的是鉴定能在凝胶图像顶部和底部产生相似强度的条带的 DNase Ⅰ 浓度。

有几种有用的方法能改进利用粗提物时的 DNase Ⅰ 足迹的结果。首先，如果在目的 DNA 元件上没有观察到保护现象，通过往结合反应中加入 2%聚乙烯醇（从 10%贮液）通常可提高分析的灵敏度。该聚合物增加提取物中蛋白质的有效浓度，可能会提高有关蛋白质-DNA 相互作用的效率。其次，如果结合反应和 DNase Ⅰ 消化都在冰上进行（除非蛋白质仅在室温条件或更高温度下与 DNA 相结合），而非在室温进行，有时可得到

一致的 DNase I 足迹实验结果。对于特定的提取物，应测试两种温度，以确定一种温度是否比另一种温度产生显著性更好的结果。当 DNase I 消化在冰上进行时，需要较高浓度的 DNase I，这是因为低温使切割率下降。对于初次进行 DNase I 足迹实验的研究人员来说，最好使用一种商业化提供的试剂盒（如 Promega 的 Core Footprinting System），这些试剂盒包含了标准试剂和阳性对照。

最初结果的解释。利用粗提物得到的 DNase I 足迹结果通常提供直接而令人信服的证据，证明提取物中序列特异性的 DNA 结合蛋白可以结合探针的特定区域。将不含 DNA 结合蛋白进行的反应与含提取物进行的反应在相邻的凝胶泳道上电泳（见方案 11.1）。为了简化对数据的解释，在整个凝胶上，除了受保护的特定区域外，对照泳道上的条带和实验泳道上的条带应具有相似的强度。如果两条泳道中总体的条带强度明显不同，可能有必要改变其中一个反应中所使用的 DNase I 的浓度。

分析粗提物时，一个另外的有用的对照是将粗提物和探针在典型结合反应中孵育，然后通过加入 $MgCl_2$ 和 $CaCl_2$，而不加入 DNase I 进行模拟消化。该反应提供了由提取物中核酸酶引起的切割产物的有关信息。提取物中的核酸酶和 DNase I 的切割位点可能不同，因此使结果的解释复杂化。如果标准结合缓冲液中也含有 $MgCl_2$，核酸酶可能会更成问题。因为大多数 DNA 结合蛋白的结合不需要 Mg^{2+}，最初的实验应该尝试利用含和不含 $MgCl_2$ 的两种结合缓冲液。如果提取物中检测到了可观的核酸酶活性，则可能有必要在足迹分析之前通过柱层析（如肝素琼脂糖层析）将提取物进行分级处理，以便将核酸酶从大多数 DNA 结合蛋白中分离出来。

除了被结合的蛋白质所保护的切割位点免受 DNase I 消化外，其他位点可能对 DNase I 切割高度敏感。经常在保护区的边界附近观察到高敏感性，这种高敏感性产生于蛋白质结合时 DNA 中的构象改变。至少有一个蛋白质家族，即 Ets 家族，会引起处于保护区中心的特有的高敏感性位点（Ernst et al. 1993)。高敏感性位点具有很大的信息量，这是因为当只有一小部分探针被蛋白质结合时，高敏感性位点就可出现，但是只有当绝大多数探针分子被蛋白质占据时，令人信服的保护才会出现。因此，没有体现保护作用的区域中高敏感性位点的存在，提示提取物中有一种蛋白质结合在附近，但是该蛋白质的浓度可能不足以形成保护。对于孤立高敏感性位点的存在，有另外两种解释：①提取物中的核酸酶可导致高敏感性切割位点的出现，因为在无提取物的对照反应中，那些位点可能不被 DNase I 切割；②当蛋白质结合远距离位点时，可能发生增强 DNase I 切割的结构改变。例如，两个蛋白质结合相互间距离远的两个位点时，两个蛋白质可能相互作用，从而引起结构改变，并增强 DNase I 对两个位点之间的序列的切割。

解释最初的 EMSA 实验时所讨论的一些问题，也与 DNase I 足迹实验相关，而其他问题则不太相关。特别是，利用 DNase I 足迹更容易确定蛋白质-DNA 相互作用的特异性。如果观察到独立的足迹，则蛋白质-DNA 相互作用对受保护区域中的序列是特异的；此种情况下，一般不需要通过竞争实验和突变体研究确定该相互作用的这种特异性。然而，为了确定这种蛋白质-DNA 相互作用是否牵扯与控制元件的功能相关的核苷酸，还必须进行竞争性和突变体实验。正如 EMSA 实验，利用放射性标记的突变体探

针，或通过利用野生型或突变型寡核苷酸的竞争实验，可以阐明这种相互作用的核苷酸要求。前一种策略一般是首选策略，这是因为该策略更不容易引起误解。如果利用相同的放射性标记引物通过 PCR 准备放射性标记的野生型和突变型探针，那么探针将具有类似的特异活性。如果某种突变扰乱了蛋白质-DNA 相互作用，那么该结果则提供有力的证据，证明被突变的核苷酸是必需的。此外，因为 DNase I 足迹反应完全在溶液中进行（即蛋白质-DNA 作用不需要在凝胶电泳步骤存活），所以通过比较放射性标记野生型和突变型探针而确定的相对亲和性与通过竞争实验确定的相对亲和性一样准确，甚至更准确。

正如 EMSA 实验所述，可用抗体确定负责 DNase I 足迹的蛋白质是否以前已被描述过。然而，只有当抗体扰乱该蛋白质的 DNA 结合活性时，它们才是有用的；抗体通常通过与 DNA 结合结构域的结合，扰乱蛋白质的 DNA 结合活性。不扰乱 DNA 结合的抗体通常对 DNase I 足迹没有显著的影响，相比之下，在 EMSA 中则观察到超迁移。

重要的是要强调 DNase I 足迹分析和可与 EMSA 联用的甲基化干扰分析提供不同类型的信息。在 DNase I 足迹分析中所观察到的受保护区域对应于蛋白质结合时 DNase I 切割不能触及的核苷酸。因为 DNase I 是具有大分子质量的蛋白质，并且 DNA 结合蛋白通常覆盖与其直接接触的核苷酸毗连的核苷酸，因此受保护区域几乎总是包括关键核苷酸两端的至少几个核苷酸。关键核苷酸通常接近受保护区域的中心，但是在缺乏突变或竞争研究的情况下，关键核苷酸难以预测。相比之下，甲基化干扰分析可提供关于蛋白质-DNA 接触的更精确的信息，这是因为在蛋白质结合之前 DNA 被小的甲基基团修饰。只有当甲基基团修饰直接接触蛋白质的鸟嘌呤或与蛋白质的僵硬区域非常近的鸟嘌呤时，甲基基团才会干扰蛋白质的结合（第 11 章）。

克隆和鉴定编码 DNA 结合蛋白基因的实验策略

上面概述了用于检测与目的 DNA 元件结合的蛋白质的最常用的策略。如果检测到了这样的活性，则可以用抗体进行实验，以确定该蛋白质是否对应于 TRANSFAC 或 JASPAR 数据库分析中预测的与该位点结合的蛋白质。也可以使用蛋白质-DNA 相互作用分析进行该蛋白质的基本表征。例如，可以分析来自一些不同细胞类型和不同发育阶段的提取物，以确定这种结合活性是否是细胞特异的或是否是受发育调控的。另外，可以分析结合位点突变体，以鉴定蛋白质-DNA 相互作用所需的核苷酸。然而，除非该蛋白质通过 TRANSFAC 或 JASPAR 搜索被成功鉴定，更加高级的研究将有赖于使用下面描述的策略之一对该蛋白质的鉴定。在该蛋白质及其编码基因被鉴定后，就可通过基因破坏或 RNA 干扰评价它们与靶基因表达的相关性，可通过诱变确定负责其活性的该蛋白质的结构域，并分析该蛋白质的作用机制和调控方式。

用于鉴定由 EMSA 或 DNase I 足迹检测到的 DNA 结合蛋白的首选方法是纯化该蛋白质并通过质谱分析来自该纯化蛋白质的蛋白质水解肽（专题 7.2）。如果目的生物体的基因组已完成测序和注释，质谱分析所得的肽的质量将可能与从一个已注释基因编码的蛋白质所预期的肽的质量相匹配，从而鉴定该蛋白质及其基因。基本的纯化策略有时是首选的，因为每个纯化步骤后，可以通过 EMSA 或 DNase I 足迹监测含有结合活性

的柱层析分步收集物。在整个纯化程序中能够跟踪结合活性可确保最终鉴定的蛋白质与最初通过 EMSA 或足迹所检测和表征的蛋白质相一致（图 7.5）。这一方法的另一优点是可以纯化依赖于两种或更多种蛋白质或亚基的结合活性；之后，可通过质谱鉴定每个亚基。相比之下，只有当结合活性为单体、同源二聚体或同源多聚体时，下文所描述的其他方法才会成功。

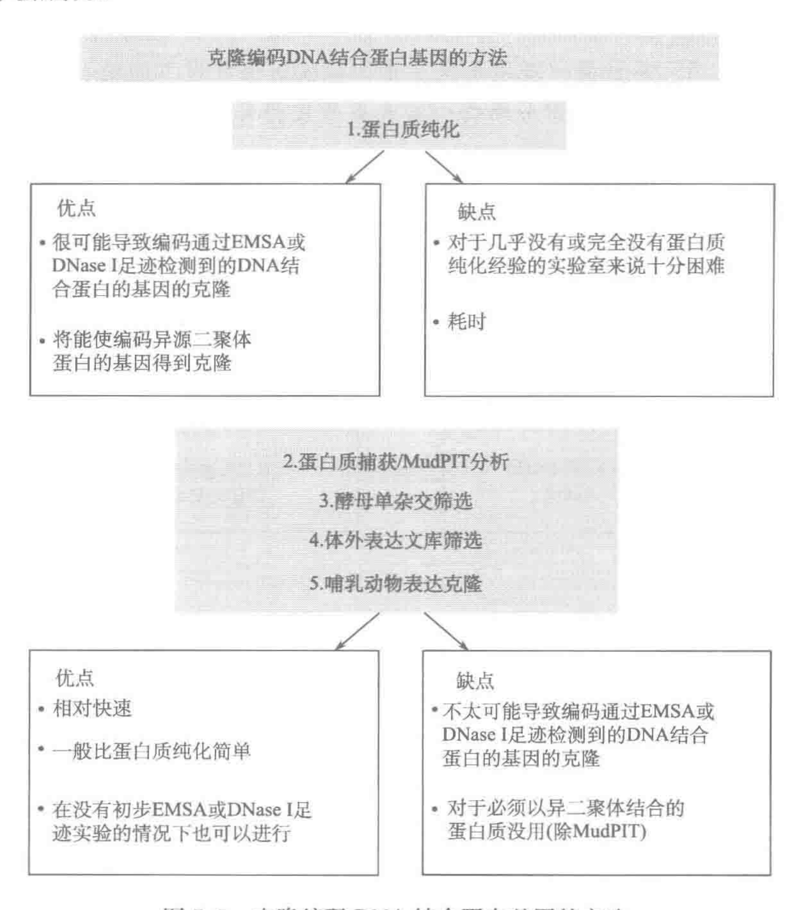


图 7.5 克隆编码 DNA 结合蛋白基因的方法。

专题 7.2 通过蛋白质纯化克隆

为了通过序列特异性 DNA 亲和层析纯化和克隆 DNA 结合蛋白（图 7.6），一般需要准备提取物丰富的细胞来源。提取物通常通过硫酸铵沉淀法浓缩，该方法也充当提取物的初步分级分离步骤。再通过一个或多个常规层析柱部分纯化浓缩的提取物，以帮助从核酸酶、蛋白酶和倾向于干扰纯化的其他 DNA 结合蛋白中分离目的蛋白。通过 EMSA 或 DNase I 足迹分析监测洗脱的层析柱分步收集物中的结合活性。收集含有大量结合活性的分步收集物，与非特异性竞争 DNA 混合，然后加到 DNA 亲和层析柱上，该

层析柱中含有共价连接目的结合位点的寡核苷酸多聚体的树脂。当样品通过树脂时，目的蛋白应与其高亲和位点结合。非特异性竞争 DNA 阻止不需要的核酸结合蛋白与树脂的稳定结合。之后，洗涤柱子，再用盐浓度逐渐增加的缓冲液洗脱结合的蛋白质。如有必要，通过第二次流过亲和柱或在含有突变型寡核苷酸序列的亲和柱上的层析，可实现进一步的纯化。通过 SDS-PAGE 检测纯化的分步收集物中的蛋白质。虽然有些方法，如变性/复性、Southwestern 印迹（如见 Gunther et al. 1990）及紫外交联（Ausubel et al. 1994，单元 12.5）等过去一直用于确定哪个凝胶条带负责目的结合活性，但成本的降低和灵敏性的增加使通过质谱分析每个候选条带成为可能。

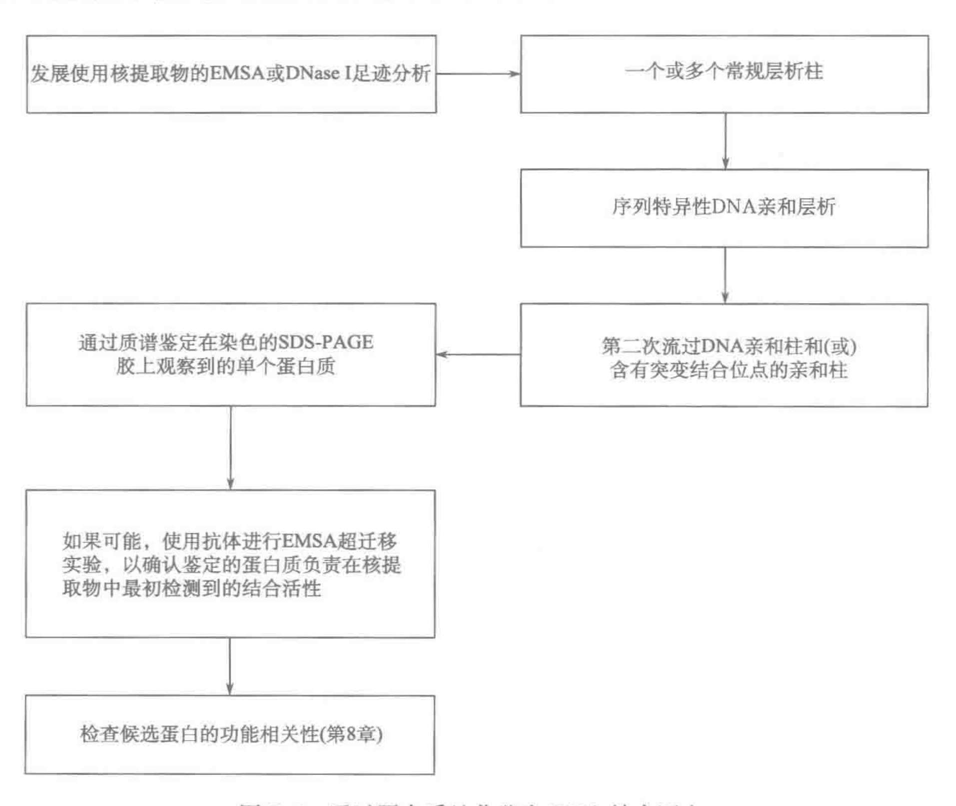


图 7.6　通过蛋白质纯化鉴定 DNA 结合蛋白。

蛋白质纯化策略的主要劣势是，即使对于有蛋白质纯化经验的实验室来说，该策略都是富有挑战性的；而对于没有蛋白质纯化经验的实验室来说，则是无法承受的。第二个局限性是有时需要大量的细胞来源。然而，随着质谱分析灵敏度的提高，成功的纯化和分析所需要的起始材料的量正在迅速下降。

与该蛋白质纯化方法相比，其他策略是容易的，但是这些策略产生编码在 EMSA 和 DNase I 足迹中观察到的蛋白质的基因的可能性较低，这些策略包括单杂交筛选、体外表达文库筛选及哺乳动物表达克隆方法（图 7.5），每一种策略概述如下。由于篇幅所

限，不包括方案，但提供恰当的参考文献。这些策略的一个共同特征是它们不依赖于蛋白质-DNA 相互作用分析方法的最初的发展。也就是说，如果在启动子或远距离控制区中已经鉴定出一个重要的 DNA 元件，那么就可以用这些技术克隆编码与该元件特异性结合的蛋白质的基因，无需对粗提物进行初步的 EMSA 或 DNase I 足迹研究。同样，通过紧随其后为 MudPIT 分析的简单的捕获实验（pulldown experiment）鉴定蛋白质的新兴策略，在缺乏一项已建立的蛋白质-DNA 相互作用分析方法的情况下也可以使用。

在缺乏预先建立的蛋白质-DNA 相互作用分析方法的情况下，DNA 结合蛋白及其编码基因的鉴定既有优势也有劣势。优势包括：如果不需要发展和优化蛋白质-DNA 相互作用分析方法，可以节省大量时间和精力；有了确定的 DNA 元件，可以直接开始克隆实验；即使初步的 EMSA 和 DNase I 足迹试验未检测到结合活性，也可以使用这些策略。

一个主要的劣势是，利用单杂交筛选或表达文库筛选分离的基因及其编码的蛋白质与目的控制元件的功能相关的可能性比较低。此外，正如引言中所述的，当使用 MudPIT 方法时，将可能鉴定出大量的候选蛋白，还需要对这些蛋白质进行评价，以鉴定出少量可能是功能相关的 DNA 序列特异性结合蛋白。

正如第 6 章中所述的，为了解释为什么单杂交筛选和表达筛选不太可能导向相关蛋白，可以考虑一下在一种功能分析中通过诱变已经仔细分析过的一个 DNA 序列元件的如下例子。如果利用粗提物检测到一个 EMSA 复合物，那么可以分析控制元件的突变体，以确定复合物的形成是否需要功能重要的核苷酸。也可以进行蛋白质表达模式的初步分析，以确定蛋白质表达模式是否与预期的模式一致。如果这些初步研究提供令人信服的证据，说明该蛋白质不是相关蛋白，则可以改变 EMSA 条件，直到鉴定出相关 DNA 结合蛋白的更有吸引力的候选蛋白。一个研究者进行这种初步表征的能力可大大提高随后纯化和克隆的蛋白质将对应于功能相关蛋白质的可能性。相比之下，如果利用单杂交筛选或表达筛选，那么在进行筛选之前，几乎没有获得任何证据，表明所鉴定的基因将编码 DNA 结合蛋白，该蛋白质识别功能相关的核苷酸并表现出预期的表达模式。只有在筛选完成后才能获得这一重要信息。

基于上述考虑，如果已经通过 EMSA 和 DNase I 足迹分析鉴定了一个相关 DNA 结合蛋白特别吸引人的候选蛋白，并且实验室具备蛋白质纯化程序的经验，则推荐蛋白质纯化策略。另外，如果蛋白质-DNA 相互作用需要异源二聚体或异源多聚体蛋白质，那么蛋白质纯化策略可能是必不可少的。此外，如果通过 EMSA 和 DNase I 足迹分析未能鉴定吸引人的候选蛋白，或实验室几乎没有蛋白质纯化的经验，则推荐将其他方法作为出发点，特别是随后为 MudPIT 分析的单杂交筛选或简单的捕获方法。

在下一节中，将描述用于克隆或鉴定编码 DNA 结合蛋白的基因的主要策略（见图 7.5），由于篇幅所限，不包括这些策略的具体方案，但提供详细方案的适当出处。

通过蛋白质纯化和肽序列分析进行克隆

用于纯化哺乳动物 DNA 结合蛋白的最成功的策略是柱层析树脂，这些树脂共价连接包含目的 DNA 序列的寡核苷酸多聚体，这种技术被称为序列特异性 DNA 亲和层析，是由 Kadonaga 和 Tjian 在纯化 Spl 蛋白时发展起来的（Kadonaga and Tjian 1986；Ka-

donaga 1991；Ausubel et al. 1994，单元 12. 10；Marshak et al. 1996）（专题 7. 2）。自那以后，这种方法已用于 100 余种蛋白质的纯化与肽序列的获取（Marshak et al. 1996）。纯化少数哺乳动物 DNA 结合蛋白没有采取 DNA 亲和层析（如见 Landschulz et al. 1988）。然而，由于细胞提取物中转录因子一般丰度低，通过常规方法纯化更困难且很少成功。在上面引用的参考文献中可找到 DNA 亲和层析方法的详细描述。想要得到 DNA 亲和层析方法的具体方案和其他建议，应查阅这些参考文献，这里仅讨论 DNA 亲和层析方法的核心问题和补充建议。

起始材料的用量

一个重要的问题是纯化哺乳动物 DNA 结合蛋白并得到其肽序列及质量信息所需要的起始材料的量。在发展出敏感的质谱分析技术之前，对于通过 Edman 降解克隆的基因，为了获得足够的肽序列信息，需要 50～100 pmol 纯化的蛋白质。由于许多序列特异性 DNA 结合蛋白在核提取物中的低丰度，以及在 DNA 亲和层析期间获得的低产量，需要 100 L 甚至更多的培养的细胞也不足为奇。此外，当需要大量起始材料时，动物组织也是有用的细胞来源。比如，NF-κB 的纯化和克隆就大大地得益于利用兔肺作为起始材料（Ghosh et al. 1990）。不幸的是，许多动物组织含有高浓度的蛋白酶。

越来越灵敏的质谱技术的出现大大减少了对大量起始材料的需求。质谱技术如今可以轻易地检测小于 1 pmol 的蛋白质。因此，对于一个纯蛋白质分析的唯一要求是能在染色的 SDS-PAGE 胶上检测到该蛋白质条带，以便从胶上切出该蛋白质条带并用于进一步的分析。即使在染色的胶上没有检测到蛋白质条带或是不知道蛋白质的分子质量，仍然可以用质谱来鉴定结合在 DNA 亲和柱上的 DNA 结合蛋白。可以通过 MudPIT 分析柱洗脱液，也可以通过 SDS-PAGE 分离洗脱液，随后将整个凝胶泳道分成 10～20 个凝胶切片，然后可以通过常规质谱分析方法（如 Nano-LC/MS/MS）分析每个切片（Domon and Aebersold 2006；Cravatt et al. 2007）。

常规层析步骤

在 DNA 亲和柱之前，通常还包括常规层析步骤。凝胶过滤通常是最有效的常规步骤（Kadonaga and Tjian 1986；Kadonaga 1991；Marshak et al. 1996），因为凝胶过滤基于与蛋白质的 DNA 结合活性无关的一个特性（即大小）来分离蛋白质。换句话说，在凝胶过滤层析中，会将目的蛋白从大多数其他核酸结合蛋白中分离出来。然而，凝胶过滤需要一个大的柱子。由于使用这样的柱子具有挑战性，因此使用含有肝素和 DEAE 树脂的离子交换柱作为第一步更普遍。虽然这些柱子使用更容易且已被成功地用于初始的分级分离（如见 Hahm et al. 1994），但是许多 DNA 结合蛋白都以类似的亲和力和这些树脂结合，因此，这些离子交换柱通常不能将目的蛋白从提取物中的大多数其他核酸结合蛋白中分离出来。亲和层析前需要的常规柱的个数随蛋白质的不同会产生变化。有些蛋白质，包括 Sp1、AP-1 和伊卡洛斯（Ikaros）蛋白，在亲和层析前仅需要进行一次凝胶过滤或离子交换柱（Kadonaga and Tjian 1986；Lee et al. 1987；Hahm et al. 1994），但是其他蛋白质，如 NF-κB，则需要数次凝胶过滤或离子交换柱（Ghosh et al. 1990）。

DNA 亲和层析

Kadonaga 和 Tjian（1986）、Kadonaga（1991）、Ausubel 等（1994，单元 12. 10）

及 Marshak 等（1996）详细描述了 DNA 亲和柱的制备和运行。一旦将蛋白质纯化到在银染凝胶上只能观察到几个高丰度条带时，可以从胶上切出每个候选条带用于质谱分析，之后紧接着评价每个候选蛋白（见专题 7.2）。

通过借助 PCR 产生多聚体可以改进 DNA 亲和层析的成功率（Hemat and McEntee 1994）。传统上，将寡核苷酸单体退火、磷酸化并连接，直到在溴化乙锭染色的琼脂糖凝胶上观察到长多聚体（Kadonaga and Tjian 1986）。虽然使用该方法已获得了相当大的成功，但由于一些不能很好解释的原因，得到长多聚体的难度相当大。通过 PCR 制备多聚体，要先合成含有结合位点二聚体的互补寡核苷酸。在每个 PCR 循环中，部分寡核苷酸以交错的方式退火，从而逐步形成长多聚体。

DNA 亲和层析程序和常规层析程序相似，然而，也有不同点，其中之一是用蛋白质过载的柱子通常可提高 DNA 亲和步骤的成功率。常规层析柱的上样量一般为 10～40 mg 蛋白质/ml，而 DNA 亲和柱的上样量有时可达 100 mg 蛋白质/ml，这样会造成过载，并使很大一部分目的蛋白将不会与柱子结合。然而，过载可使柱子被目的蛋白所饱和，这会增加高盐洗脱物中蛋白质的浓度。相对较高的浓度有利于蛋白质的检测，并会增加蛋白质的稳定性。虽然过载大大减少了纯蛋白的总产量，但为了通过质谱分析蛋白质，可能仍有必要对蛋白质进行纯化，使其达到足够的纯度。也就是说，可能需要牺牲产量以提高纯度。

常规层析和 DNA 亲和柱之间的第二个不同点是，小规模（即试验性的）的 DNA 亲和柱通常难以成功。当发展常规层析程序时，通过在小型柱上载样相对少量的蛋白质，确定层析条件是很常见的，在某些情况下蛋白质上样量甚至小于每毫升树脂的建议用量。而对于 DNA 亲和柱，这样做通常不会成功，因为洗脱物中的蛋白质可能不稳定或浓度不足以检测。如果在小规模的试点实验中未得到结合活性，那么可能需要在更大的规模上重复实验。

向 DNA 亲和柱中加入非特异竞争性 DNA 也可以牺牲总产量以获得纯度。通常在上柱前将竞争性 DNA 和蛋白质样品进行孵育。Kadonaga（1991）和 Marshak 等（1996）描述了确定竞争性 DNA 加入量的程序。竞争性 DNA 和那些不能和连在柱上的寡核苷酸特异性结合的 DNA 结合蛋白相结合，从而阻止这些 DNA 结合蛋白与柱的高效结合。实际上，竞争性 DNA 也在某种程度上组织目的蛋白和柱子的结合。有时可以通过减少竞争性 DNA 的用量来增大目的蛋白的产量。但是，为了得到目的蛋白的最佳纯度，可能需要高浓度的竞争性 DNA。

已鉴定基因编码目的 DNA 结合活性的确证

为了确定由 DNA 亲和层析和质谱鉴定的基因编码最初由 EMSA 或 DNase I 足迹分析表征的蛋白质，需要两种实验策略。第一种方法是制备重组蛋白（通过体外翻译或在 *E. coli*、HEK293T 细胞中表达），并将重组蛋白的 DNA 结合特征与粗提物中检测到的蛋白质的 DNA 结合特征相比较。重组蛋白应和粗提物中纯化的蛋白质结合一样，结合具有相同核苷酸要求的相同 DNA 元件。而且，如果重组蛋白是全长的，它可能产生与粗提物中检测到的复合物共迁移的 EMSA 复合物，或可能产生一种 DNase I 足迹或甲基化干扰模式，该模式与粗制或纯化蛋白质所产生的模式难以区分。第二种方法是制备

或获得针对重组蛋白或人工合成多肽的抗体，并表明它们和粗制及纯化蛋白质发生反应。抗体应该能超迁移或破坏用粗提物所观察到的 EMSA 复合物。如果这些实验确认该基因编码目的蛋白质，则可以使用第 8 章所描述的方法来评定该蛋白质在它结合的控制元件的功能中的潜在作用。

其他克隆方法

单杂交筛选

一般在酵母中进行的单杂交筛选，有时是一种用于鉴定编码与目的 DNA 元件结合的蛋白质的基因的强有力的方法（Fields and Song 1989；Li and Herskowitz 1993；Wang and Reed 1993；Inouye et al. 1994）。首先，构建报告质粒，该质粒包含 TATA 框上游多拷贝目的结合位点和报告基因（如 HIS3 基因或 *lacZ* 基因）（图 7.7）。此报告质粒可稳定地整合进酵母基因组中。然后，准备含有来源于合适细胞的 cDNA 的文库。文库的设计要使编码的蛋白质以融合蛋白的形式合成，并包含强转录激活结构域。用文

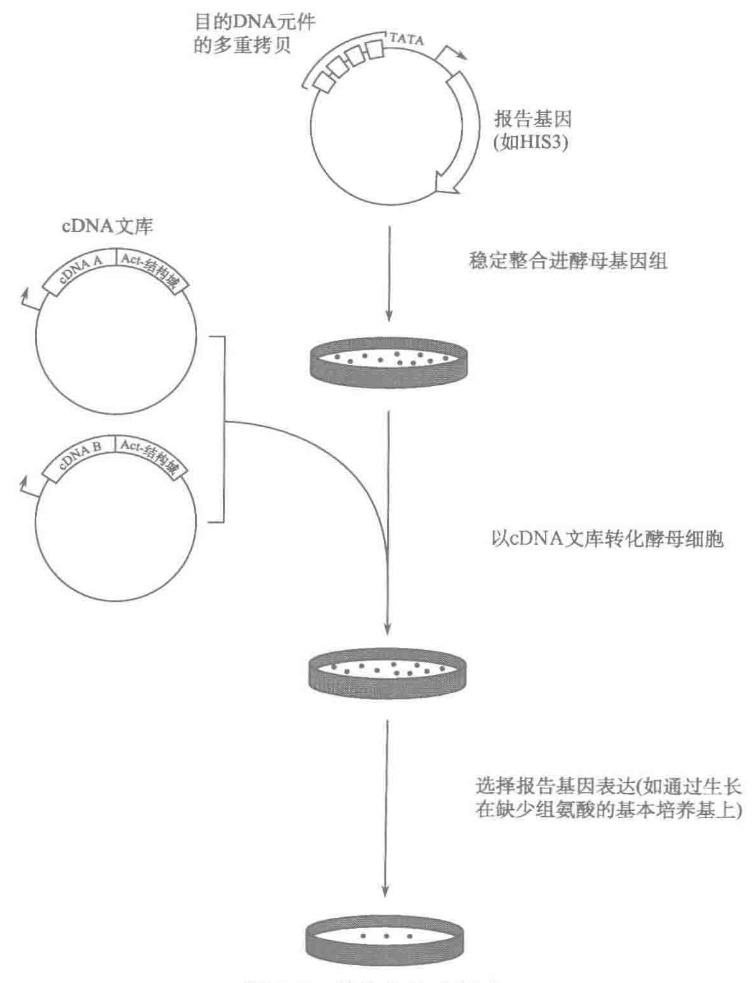


图 7.7 单杂交筛选策略。

库转化包含报告基因的酵母细胞。表达可与多聚化位点结合的融合蛋白的酵母细胞也表达报告基因，并且报告基因的表达通常采用选择策略进行监测。例如，通过在缺少组氨酸的基本培养基上生长细胞克隆，监测 HIS3 报告基因的表达。最后，可从经过选择的细胞中分离 cDNA，并进一步表征以确证编码的蛋白质与目的 DNA 序列结合。

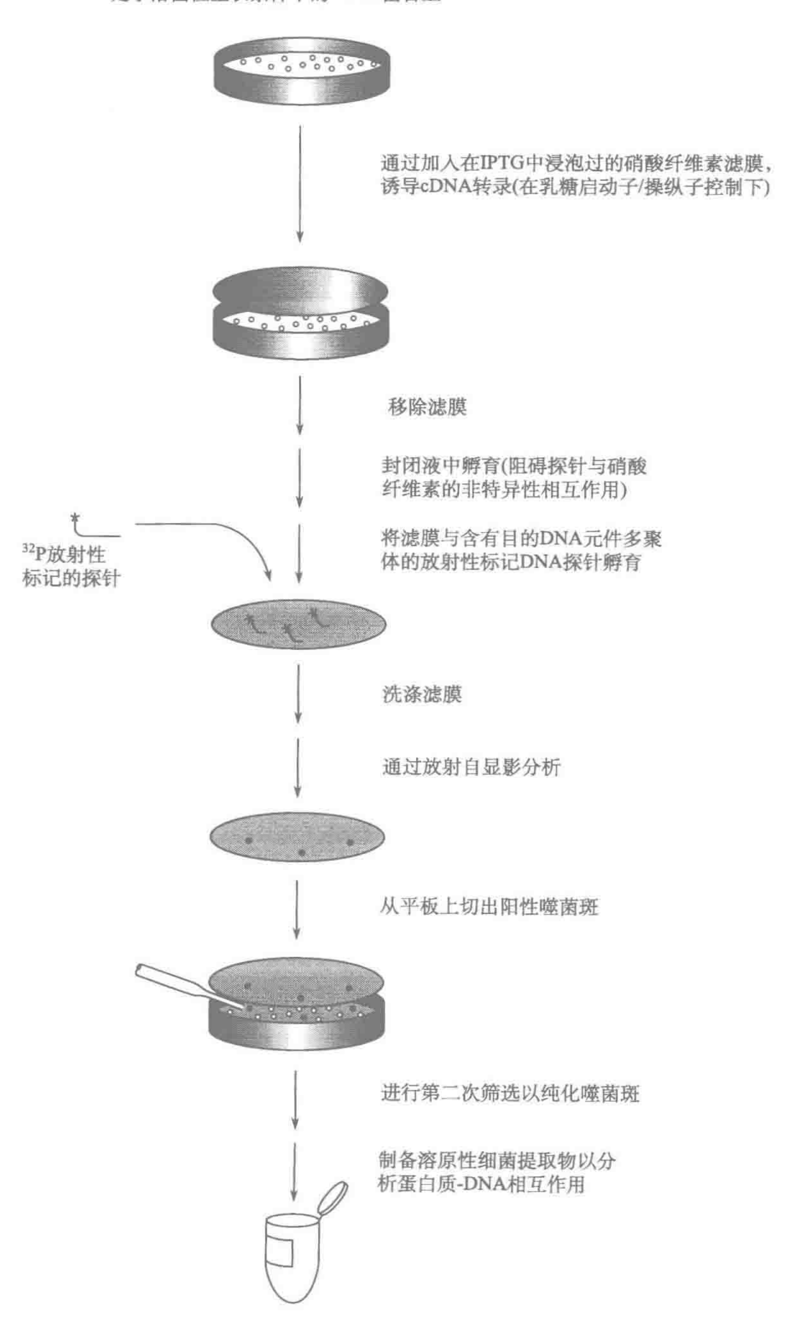
单杂交法已被用于克隆数个重要 DNA 结合蛋白，并且该方法具有数个优点。第一，该方法相对简单和快速。第二，蛋白质因在相对天然的体内条件下结合而被筛选。相比之下，蛋白质纯化和表达文库筛选（下文描述）则依赖于体外结合。虽然在酵母中表达的哺乳动物蛋白质不太可能获得在哺乳动物细胞中发现的所有翻译后修饰，但它们仍可能获得一处或多处重要的修饰。最后一个优点是该方法通常非常灵敏，并能鉴定仅以中等亲和力结合的蛋白质。

这些优点需要与两个显著的缺点平衡，包括：① 只有当结合需要蛋白质单体、同源二聚体或同源多聚体时，该方法一般才会成功；② 该程序不一定能鉴定编码一个来自经 EMSA 或 DNase I 足迹分析检测的哺乳动物粗提物中的蛋白质的基因。单杂交方法的另一需要注意的是所用的结合位点必须不能与内源性酵母激活因子蛋白相互作用。如果酵母激活因子与该位点相结合，则在相关 cDNA 表达质粒不存在时也会观察到报告基因的表达。为了克服这一问题，可能使用结合位点的变体，该变体在其天然环境中保持活性但不与酵母激活因子结合。

尽管该方法存在上述缺点，但如果有蛋白质单体、同源二聚体或同源多聚体可识别该位点的合理的机会，强烈建议将单杂交程序作为基因鉴定的出发点。进行单杂交筛选详细的程序可从上面引用的论文中找到，也可以从 Clontech 公司获得，该公司出售进行单杂交筛选和制备单杂交文库（the Matchmaker One-Hybrid Library Construction & Screening Kit）的试剂盒。值得一提的是，酵母单杂交分析已经被修改用于可与一系列启动子结合的蛋白质的高通量鉴定，以便阐明转录调控网络。

利用 DNA 或抗体作为探针进行体外表达文库筛选

在发展出单杂交筛选方法之前，用于克隆编码新的 DNA 结合蛋白的基因的常用方法涉及利用放射性标记的 DNA 作探针进行表达文库的筛选，该方法（Singh et al. 1988）是基因克隆技术的重要进展。关于表达文库筛选方法的详细描述，见 Ausubel 等（1994，单元 12.7）的研究。该技术需要以来自目的细胞类型的 mRNA 制备的 cDNA 文库，文库必须构建在 λ 噬菌体载体中，以保证感染 *E. coli* 后插入 cDNA 编码的蛋白质可诱导表达。建议使用 λgt11 作为载体，该载体表达的蛋白质是 β-半乳糖苷酶融合蛋白。利用 oligo（dT）或随机寡核苷酸为引物构建文库，已获得一些成功的结果。将噬菌体文库铺在处于裂解生长条件下的细菌平板上（图 7.8）。当噬菌斑出现时，将诱导剂 IPTG 浸泡过的硝酸纤维素膜铺在平板上，诱导 cDNA 转录。当噬菌体裂解时，被诱导的蛋白质黏附到硝酸纤维素膜上。从平板上取下滤膜，放入封闭液中孵育，以防止探针和硝酸纤维素滤膜之间的非特异性相互所用。然后，用含有多拷贝目的结合位点的放射性标记 DNA 片段探测滤膜。探针应与含有序列特异性 DNA 结合蛋白的噬菌斑特异性结合。在结合反应体系中，加入非特异性竞争 DNA 以抑制放射性标记探针产生非特异性相互作用。在有些程序（Vinson et al. 1988）中，在滤膜上表达的蛋白质首先经过



用cDNA表达文库感染*E. coli*，并铺在处于溶菌性生长条件下的*E. coli*菌苔上
通过加入在IPTG中浸泡过的硝酸纤维素滤膜，诱导cDNA转录(在乳糖启动子/操纵子控制下)
移除滤膜
封闭液中孵育(阻碍探针与硝酸纤维素的非特异性相互作用)
^{32}P放射性标记的探针
将滤膜与含有目的DNA元件多聚体的放射性标记DNA探针孵育
洗涤滤膜
通过放射自显影分析
从平板上切出阳性噬菌斑
进行第二次筛选以纯化噬菌斑
制备溶原性细菌提取物以分析蛋白质-DNA相互作用

图 7.8 体外表达克隆。

变性/复性处理，以达到增加能与探针相结合的正确折叠的蛋白质分子数的目的。分离阳性噬菌斑并重新筛选，以纯化噬菌体，然后制备溶源性细菌提取物，通过 EMSA 表征重组的融合蛋白。

与单杂交方法类似，如果蛋白质必须以异聚体的形式与 DNA 相结合，表达文库筛选方法就难以成功，而且不能鉴定编码 EMSA 或 DNase I 足迹分析中检测到的蛋白质的基因。另外，产生比活性足够高和结合位点拷贝数足够多的探针分子可能具有挑战性。在最初的方案中，探针的制备是通过从合适的质粒中切出含有多拷贝结合位点的限制性酶切片段，再进行磷酸化（Singh et al. 1988）。使用 PCR 制备长的多聚探针（如上所述制备 DNA 亲和柱），可提高这种方法的成功率（Hemat and McEntee 1994；Schmitt and McEntee 1996）。

对于大多数研究来说，单杂交筛选可能比体外表达文库筛选更具有吸引力。因为单杂交筛选不需要进行耗时的噬菌体滴度步骤，而这些步骤在进行体外表达文库筛选前一般是必需的。另外，单杂交筛选不需要大量的放射性元素。最后，单杂交筛选对体外筛选期间可遇到的蛋白质折叠和稳定性问题不太敏感。因此，如果单杂交筛选和蛋白质纯化策略无法鉴定相关性转录因子，首先推荐体外表达文库筛选。

哺乳动物表达克隆方法

哺乳动物表达克隆法用于新基因鉴定已有 30 年的历史。基本上，通常通过转染剪切的基因组 DNA 或 cDNA 表达文库，将目的基因转移到受体哺乳动物细胞系中，接受目的基因的细胞通过选择或筛选进行鉴定。

使用表达克隆选择策略的一个早期的例子是基于原癌基因赋予受体细胞转化表型能力的原癌基因的鉴定（如见 Shih et al. 1979）。筛选策略已经用于分离很多新基因，包括编码细胞表面蛋白质的几个基因。一个常见的筛选程序需要针对目的蛋白的抗体及在质粒载体中构建 cDNA 表达文库（Seed 1987；Ausubel et al. 1994，单元 6.11）。表达文库包含位于强启动子和增强子下游的 cDNA。质粒上还含有 SV40 复制起点，该复制起点使质粒在表达 SV40 T 抗原的 COS-7 猴肾细胞中以高拷贝数复制。将文库分为几个池，每个池中含有不同克隆的混合物。之后，向各池中引入 COS-7 细胞，从而导致质粒复制和 cDNA 表达。为分离得到编码细胞表面蛋白的基因，通过淘选分析每个转染细胞群的蛋白质表达，这涉及将在塑料平皿上固定特异性抗体，从而促使表达表面蛋白的细胞的黏附，以及非表达细胞的除去。在 *E. coli* 中分离和扩增来自黏附细胞的质粒 DNA，所得到的质粒再一次分成不同的池，并且重复整个程序（即转染、黏着斑筛选、质粒回收和扩增）数次，最后导向含有目的 cDNA 的单一质粒类型。

由于上述哺乳动物表达克隆策略的成功，几个实验室已尝试利用相似的程序克隆与特定调控元件或调控区相结合的转录因子。对于某些实验，用含有多拷贝目的位点的合成启动子驱动显性选择标记基因的表达。在其他实验中，将编码细胞表面蛋白质的基因置于合成启动子的下游，以期利用上述黏着斑筛选程序分离编码相关转录因子的 cDNA。也检验了天然的组织特异性启动子，假定受体细胞含有转录所需的所有遍在因子，当提供编码必要组织特异性因子的 cDNA 时，使启动子能发挥功能。不幸的是这些策略很少成功，它们的失败可能是综合因素的结果，特别是单一转录因子无力诱导转

录的水平达到足够高出背景的程度。

尽管存在这些缺点，哺乳动物表达克隆法偶尔被用来克隆编码对基因调控重要的蛋白质的基因。其中的一个例子是克隆编码 SCAP 的基因（SREBP 切割活化蛋白；Hua et al. 1996），SCAP 调控甾醇应答元件结合蛋白 SREBP-1 的功能（Yokoyama et al. 1993）。为了克隆 SCAP，质粒 cDNA 文库，连同含有在多重 SREBP-1 结合位点控制下的萤光素酶报告基因的质粒，一起转化每池 1000 个克隆的人 293 细胞。文库利用从甾醇抗性细胞中提取的 mRNA，来分离负责甾醇抗性的基因。作为对甾醇超载的应答，正常细胞下调 SREBP-1 应答基因的表达。因此，表达赋予甾醇抗性的蛋白质的细胞，在高浓度甾醇存在的条件下，应表现出比正常的 293 细胞更高的萤光素酶活性。的确，经鉴定，质粒池产生高萤光素酶活性。质粒子池的顺序分析导致 SCAP 基因的克隆。SCAP 的功能分析表明它切割膜结合的 SREBP-1，使其转移到细胞核中，并激活转录。该例子产生了一种可能性，即用于制备和分析质粒文库的技术的改进将允许通过相似的方法克隆其他转录调控因子。

参考文献

Ausubel F. M., Brent R. E., Kingston E., Moore D. D., Seidman J. G., Smith J. A., and Struhl K. 1994. *Current protocols in molecular biology*. John Wiley and Sons, New York.

Cravatt B. F., Simon G. M., and Yates J. R., 3rd. 2007. The biological impact of mass-spectrometry-based proteomics. *Nature* 450: 991-1000.

Deplancke B., Dupuy D., Vidal M., and Walhout A. J. M. 2004. A gateway-compatible yeast one-hybrid screen. *Genome Res*. 14: 2093-2101.

Deplancke B., Vermeirssen V., Arda H. E., Martinez N. J., and Walhout A. J. M. 2006. Gateway-compatible yeast one-hybrid screens. *CSH Protocols* doi: 10. 1101/pdb. prot4590.

Dignam J. D., Lebovitz R. M., and Roeder R. G. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res*. 11: 1475-1489.

Domon B. and Aebersold R. 2006. Mass spectrometry and protein analysis. *Science* 312: 212-217.

Dynan W. S. and Tjian R. 1983. The promoter-specific transcription factor Spl binds to upstream sequences in the SV40 early promoter. *Cell* 35: 79-87.

Emami K. H., Jam A., and Smale S. T. 1997. Mechanism of synergy between TATA and initiator: Synergistic binding of TFIID following a putative TFIIA-induced isomerization. *Genes Dev*. 11: 3007-3019.

Ernst P., Hahm K., and Smale S. T. 1993. Both LyF-1 and an Ets protein interact with a critical promoter element in the murine terminal transferase gene. *Mol. Cell. Biol*. 13: 2982-2992.

Fields S. and Song O. 1989. A novel genetic system to detect protein-protein interactions. *Nature* 340: 245-246.

Fournier M. L., Gilmore J. M., Martin-Brown S. A., and Washburn M. P. 2007. Multidimensional separations-based shotgun proteomics. *Chem. Rev*. 107: 3654-3686.

Ghosh S., Gifford A. M., Riviere L. R., Tempst P., Nolan G. P., and Baltimore D. 1990. Cloning of the p50 DNA binding subunit of NF-κB: Homology to rel and dorsal. *Cell* 62: 1019-1029.

Gorski K., Carneiro M., and Schibler U. 1986. Tissue-specific in vitro transcription from the mouse albumin promoter. *Cell* 47: 767-776.

Gunther C. V., Nyc J. A., Bryner R. S., and Graves B. J. 1990. Sequence-specific DNA binding of the proto-oncoprotein Ets-1 defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus. *Genes Dev*. 4: 667-678.

Hahm K., Ernst P., Lo K., Kim G. S., Turck C., and Smale S. T. 1994. The lymphoid transcription factor LyF-1 is encoded by specific, alternatively spliced mRNAs derived from the Ikaros gene. *Mol. Cell. Biol.* 14: 7111-7123.

Harlow E. and Lane D. 1999. *Using antibodies: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Hazbun T. R. and Fields S. 2002. A genome-wide screen for site-specific DNA-binding proteins. *Mol. Cell. Proteomics* 1: 538-543.

Hemat F. and McEntee K. 1994. A rapid and efficient PCR-based method for synthesizing high-molecular-weight multimers of oligonucleotides. *Biochem. Biophys. Res. Comm.* 205: 475-481.

Hua X., Nohturfft A., Goldstein J. L., and Brown M. S. 1996. Sterol resistance in CHO cells traced to point mutation in SREBP cleavage-activating protein. *Cell* 87: 415-426.

Inouye C., Remondelli P., Karin M., and Elledge S. 1994. Isolation of a cDNA encoding a metal response element binding protein using a novel expression cloning procedure: The one hybrid system. *DNA Cell Biol.* 13: 731-742.

Kadonaga J. T. 1991. Purification of sequence-specific binding proteins by DNA affinity chromatography. *Methods Enzymol.* 208: 10-23.

Kadonaga J. T. and Tjian R. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci.* 83: 5889-5893.

Landschulz W. H., Johnson P. F., Adashi E. Y., Graves B. J., and McKnight S. L. 1988. Isolation of a recombinant copy of the gene encoding C/EBP. *Genes Dev.* 2: 786-800.

Lee W., Mitchell P., and Tjian R. 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. *Cell* 49: 741-752.

Li J. J. and Herskowitz I. 1993. Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. *Science* 262: 1870-1874.

Lo K., Landau N. R., and Smale S. T. 1991. LyF-1, a transcriptional regulator that interacts with a novel class of promoters for lymphocyte-specific genes. *Mol. Cell. Biol.* 11: 5229-5243.

Marshak D. R., Kadonaga I. T., Burgess R. R., Knuth M. W., Brennan W. A., Jr., and Lin S. -H. 1996. *Strategies for protein purification and characterization: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Pierce J. W., Lenardo M., and Baltimore D. 1988. Oligonucleotide that binds nuclear factor NF-κB acts as a lymphoid-specific and inducible enhancer element. *Proc. Natl. Acad. Sci.* 85: 1482-1486.

Reed D. E., Huang X. M., Wohischlegel J. A., Levine M. S., and Senger K. 2008. DEAF-1 regulates immunity gene expression in *Drosophila*. *Proc. Natl. Acad. Sci.* 105: 8351-8356.

Sambrook J. and Russell D. W. 2001. *Molecular cloning: A laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schmitt A. P. and McEntee K. 1996. Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* 93: 5777-5782.

Seed B. 1987. An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. *Nature* 329: 840-842.

Shapiro D. J., Sharp PA., Wahli W. W., and Keller M. J. 1988. A high-efficiency HeLa cell nuclear transcription extract. *DNA* 7: 47-55.

Shih C., Shilo B. Z., Goldfarb M. P., Dannenberg A., and Weinberg R. A. 1979. Passage of phenotypes of chemically transformed cells via transfection of DNA and chromatin. *Proc. Natl. Acad. Sci.* 76: 5714-5718.

Singh H., LeBowitz J. H., Baldwin A. S., Jr., and Sharp P. A. 1988. Molecular cloning of an enhancer binding protein: Isolation by screening of an expression library with a recognition site DNA. Cell 52: 415-423.

Soeller W. C., Poole S. J., and Kornberg T. 1988. In vitro transcription of the *Drosophila* engrailed gene. *Genes Dev*. 2: 68-81.

Vinson C. R., LaMarco K. L., Johnson P. F., Landschulz W. H., and McKnight S. L. 1988. In Situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. *Genes Dev*. 2: 801-806.

Wang M. M. and Reed R. R. 1993. Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. *Nature* 364: 121-126.

Washburn M. P., Wolters D., and Yates J. R., 3rd. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol*. 19: 242-247.

Yokoyama C., Wang X., Briggs M. R., Admon A., Wu J., Hua X., Goldstein J. L., and Brown M. S. 1993. SREBP-1, a basic-helixloop-helix-leucine zipper protein that controls transcription of the low density lipoprotein receptor gene. *Cell* 75: 187-197.

网络资源

http://jaspar.genereg.net/：JASPAR 数据库是一个经实验界定的多细胞真核生物转录因子结合位点的审核网站。

http://www.gene-regulation.comlpub/databases.html：TRANSFAC 数据库含有转录因子、转录因子靶位点和转录因子调控基因的相关数据。

8 确认蛋白质-DNA 相互作用的功能相关性

要点

- 蛋白质-DNA 相互作用的功能相关性难以确定。
- 可以通过几种不同的实验测验相互作用是功能相关的假设，但没有一种实验独自达成结论。

引言和概述

在现代分子生物学时代，鉴定与目的控制区相互作用的 DNA 结合蛋白是相对简单的，但确切地确定特定转录因子通过与界定的控制元件相结合而直接调控靶基因却是最困难的工作。前面各章中描述了鉴定重要 DNA 序列元件及其结合蛋白的实验策略。在大多数情况下，通过 TRANSFAC 和 JASPAR 数据库搜索、与元件体外结合的蛋白质的大规模质谱分析、转录因子结合位点的全基因组染色质免疫沉淀分析（见第 1 章和第 9 章）或酵母单杂交筛选，转录因子总会与潜在的基因调控因子发生牵连。候选 DNA 结合蛋白的鉴定是一个很大的进步，因为它使人们推测这种蛋白质在邻近基因的调控中担负控制元件的功能。但是，单独蛋白质的鉴定并不能证明其生物相关性。例如，检测细胞核提取物中的蛋白质-DNA 相互作用反映出许多因素，包括：①制备提取物的细胞内蛋白质的丰度；②蛋白质从细胞中提取的效率；③提取物中活性蛋白的稳定性；④提取物制备期间重要翻译后修饰的维持；⑤用于体外 DNA 结合分析的条件；⑥蛋白质对分离的控制元件的亲和力（表 8.1）。

表 8.1 影响蛋白质-DNA 相互作用体外检测和体内相关性的因素

体外检测	体内功能相关性
· 细胞内蛋白质的丰度	· 细胞核中蛋白质的丰度和稳定性
· 蛋白质提取的效率	· 蛋白质对位点的亲和力
· 抽提物中蛋白质的稳定性	· 蛋白质与其他蛋白质适当互作的能力
· 抽提物制备期间重要翻译后修饰的维持	· 使蛋白质进行必要互作的翻译后修饰
· 用于体外 DNA 结合分析的条件	· 蛋白质的合适亚细胞定位和亚核定位
· 蛋白质对分离的控制元件的亲和力	

检测体外蛋白质-DNA 相互作用的标准与确定哪种蛋白质在体内与控制元件发生功能性相互作用（即哪种蛋白质通过与控制元件的结合调控内源基因）（表 8.1）的标准非常不同。这些标准包括：①细胞核中蛋白质的丰度和稳定性；②蛋白质对位点的亲和力；③蛋白质与结合于邻近位点的其他蛋白质及非 DNA 结合性辅因子发生合适相互作用的能力；④使蛋白质进行必要的蛋白质-DNA 和蛋白质-蛋白质相互作用的适当的翻译后修饰；⑤蛋白质合适的亚核定位。

当考虑到以上各点时，显而易见的是在体外检测蛋白质-DNA 相互作用仅为其体内相关性提供不充分的证据，即使已知蛋白结合的 DNA 序列元件是重要的。若能够识别 DNA 序列元件的蛋白质是在 TRANSFAC 或 JASPAR 数据库搜索中被鉴定的，也存在类似水平的不确定性，因为用于创建这些数据库的一致识别序列是运用体外蛋白质-DNA 相互作用分析确定的。如第 7 章中所讨论的，大多数 DNA 结合蛋白能以广泛的亲和力识别广谱 DNA 序列。此外，大多数 DNA 结合蛋白是多蛋白家族的成员，同时每种细胞类型都包含几个识别相似 DNA 序列的家族成员。基于这些考虑，很可能多种蛋白质将能在体外结合确定的控制元件，包括特定蛋白家族的几个成员，以及可能识别相似或重叠序列的另一家族的成员。目前的挑战是确定这些蛋白质中的哪个蛋白质能进行蛋白质-蛋白质和蛋白质-DNA 相互作用，并借此调控内源基因。如果碰巧在体外只检测到一种占优势的 DNA 结合蛋白，或数据库搜索仅鉴定出一种候选蛋白，该蛋白质仍可能不是负责体内控制元件功能的 DNA 结合蛋白。细胞中的其他蛋白质几乎肯定能够结合相同的元件，即使在最初的数据库分析或电泳迁移率变动分析（EMSA）研究中没有检测到它们。

还没有发展出一种证实哺乳动物细胞内蛋白质-DNA 相互作用的功能重要性的决定性的方法。然而，两种强大的技术使评价与目的 DNA 元件相互作用的 DNA 结合蛋白的功能相关性的努力发生了革命性的变化，即 ChIP 和 RNA 干扰（RNAi）。虽然这两种技术在本书第一版（2000 年出版）前都已产生，但直到 21 世纪，这两种技术才成为研究真核细胞，特别是哺乳动物细胞中蛋白质-DNA 相互作用的常用方法。

如下文和第 9 章所述，ChIP 技术现在已经成为监测活细胞中蛋白质与特定基因组区域结合的非常常用的方法。在控制良好的 ChIP 实验中，阳性结果提供了之前无法获得的至关重要的信息，即候选蛋白确实结合在目的控制元件附近，这为候选 DNA 结合蛋白负责 DNA 控制元件的功能的假设提供了有力的支持。然而，ChIP 并不是一种功

能性分析，且仅仅显示体内的结合。因为数项研究表明，DNA 结合蛋白并不是在它们体内结合的所有 DNA 元件上都有功能活性（见 Walter et al. 1994；Li and Johnston 2001；Martone et al. 2003；Phuc Le et al. 2005；Beima et al. 2006；Hollenhorst et al. 2007；Dong et al. 2008），必须获得功能相关性的其他证据。

DNA 结合蛋白的功能缺失研究对于评价蛋白质-DNA 相互作用的功能重要性至关重要。然而，在 RNAi 成为常用分子生物学工具之前，可用于进行这类研究的方法，特别是在哺乳动物细胞中，要么费力（如通过同源重组进行基因破坏），要么不稳定，因为它们的成功率相对较低（如反义 RNA 或显性失活方法）。现在可以利用 RNAi 迅速敲低编码几乎任何蛋白质的基因的表达，只要获得 RNAi 易转染的细胞群。遗憾的是，尽管 RNAi 功能强大，但它也不能轻易克服所有功能缺失方法固有的局限性，包括两个或两个以上因子间的冗余造成的困难，以及辨别直接和间接效应的困难等。

再次声明，虽然 ChIP 和 RNAi 已经成为评价蛋白质-DNA 相互作用的相关性的两种最有用的方法，但两者都不能产生决定性的结果，并且都具有很大的局限性。在没有决定性的实验（也就是让研究者直接观察到蛋白质结合处于自然染色体位置内的控制元件，并且调控与该元件连锁的基因的转录的实验）的情况下，唯一可行的方法是假设蛋白质-DNA 相互作用是相关的，然后使该假设接受尽可能多的严格的测验，其中，ChIP 和 RNAi 实验就充当着两种比较常见且重要的测验。至于在什么程度上就可以肯定蛋白质-DNA 相互作用对调控目的基因具有功能重要性的假设已经被确证，则取决于主观评价，并且因此最好留给科学界判断。

本章中介绍了 12 种不同的方法，首先是 ChIP 和功能缺失方法，它们可用于测验特定的蛋白质-DNA 相互作用是否具有功能重要性。我们描述了从每种方法获得的信息并解释了为什么每种方法产生有用的但非决定性的结果，这些方法在所需的工作量和所获得的信息的质量方面千差万别。如果主要目的是确立相互作用的功能重要性，就应该采取几种不同的方法。

重要的是要强调严格的检验转录因子-DNA 相互作用的相关性可能不是研究者的重要目标。主要的兴趣可能反而是确定一个因子对于生物过程是否重要，该因子是否是在检验一种特定基因的调控的过程中发现的。例如，即使没有该转录因子调控目的基因的强大证据，它本身可能有一种有趣的表达模式，该模式值得进一步探索其生物功能。在这种情况下，下一步更明显：破坏细胞或动物中编码这种因子的基因，或通过 RNAi 减少基因表达。然后可以进行敲除或敲低表型的全面表征，包括鉴定因子缺乏时错误表达的基因的微阵列实验。如果观察到有趣的表型，研究者可能对该因子开始一个有广泛基础的分析，不管它是否与起初用于鉴定它的控制元件发生功能性相互作用，这是常见的且有效的行动方针。然而，如果长期目标是要把蛋白质和与它相关的靶基因联系起来，或者要仔细分析 DNA 结合蛋白质、辅因子及通用转录因子联合诱导靶基因的机制，则最终必须考虑本章中所讨论的问题。

实验策略

染色质免疫沉淀

通常测验特定的 DNA 结合蛋白是否负责控制区的功能的最重要的实验之一是 ChIP 分析（图 8.1；实验方案 9.3）。Gilmour 和 Lis 发展出 ChIP 分析的先驱，它依赖于紫外交联和免疫沉淀的 DNA 片段的 Southern 印迹或斑点印迹检测，以此来监测大肠杆菌和果蝇中 RNA 聚合酶结合的转录基因和蓄势基因（poised gene）（Gilmour and Lis 1984，1985，1986）。当代 ChIP 分析，通常依赖于蛋白质到 DNA 的甲醛交联（Solomon and Varshavsky 1985；Solomon et al. 1988；Braunstein et al. 1993；Orlando and Paro 1993；Orlando et al. 1997）及 DNA 片段的 PCR（聚合酶链反应）检测，该方法首先被 Grunstein 及其同事用于监测酿酒酵母（*Saccharomyces cerevisiae*）中的组蛋白修饰和 SIR 蛋白质结合（见实验方案 9.3；Hecht et al. 1996；Rundlett et al. 1998；Hecht and Grunstein 1999）。ChIP 程序很快被改造用于哺乳动物细胞，最早用紫外交联，随后用甲醛交联（Boyd and Farnham 1997；Boyd et al. 1998；Wathelet et al. 1998；Parekh and Maniatis 1999）。

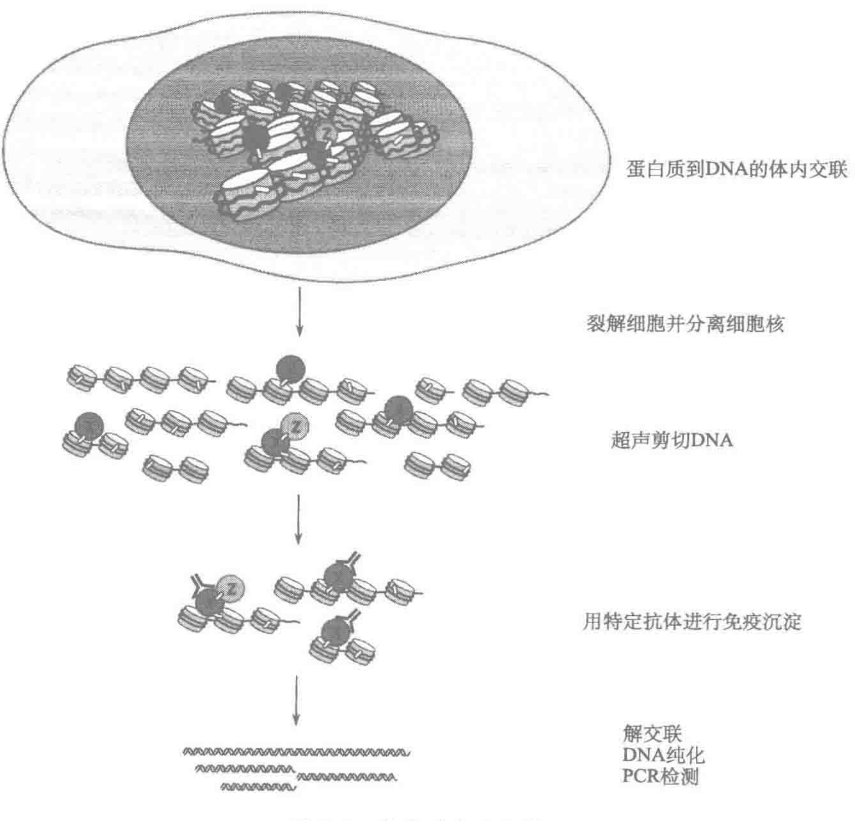


图 8.1 染色质免疫沉淀。

简而言之，用甲醛处理生长的细胞，将 DNA 结合蛋白交联到它们的靶位点上（方案 9.3），然后裂解这些细胞，并通过超声波处理和（或）核酸酶消化将 DNA 切割成片段。以抗目的 DNA 结合蛋白的抗体进行免疫沉淀，纯化蛋白质-DNA 复合物。为了确定蛋白质是否与假定的靶元件交联，将免疫沉淀产物解交联，并且通过常规或实时 PCR 监测目的 DNA 片段的存在。

Maniatis 实验室是最早在哺乳动物细胞中使用该程序的实验室之一，他们将该程序用于分析 β-干扰素（IFN-β）增强子（图 8.2）（Wathelet et al. 1998）。Maniatis 实验室进行的其他实验涉及两种干扰素调节因子（IRF）家族成员——IRF-3 和 IRF-7，并将其作为 IFN-β 增强子的相关激活因子，IFN-β 增强子由病毒感染诱导。为进一步测验这个假设，研究者用模拟（mock）感染和病毒感染的细胞进行 ChIP 分析。在甲醛交联、提取物制备和 DNA 片段化之后，使用抗各种 IRF 家族成员的抗体进行免疫沉淀反应。用 PCR 分析免疫沉淀的 DNA，结果揭示 IFN-β 增强子片段存在于用抗 IRF-3 和 IRF-7 而不是 IRF-1 抗体获得的免疫沉淀中（图 8.2）。重要的是，IFN-β 增强子片段仅存在于来自病毒感染的细胞而不是模拟感染的细胞的免疫沉淀中。使用抗其他 DNA 结合蛋白

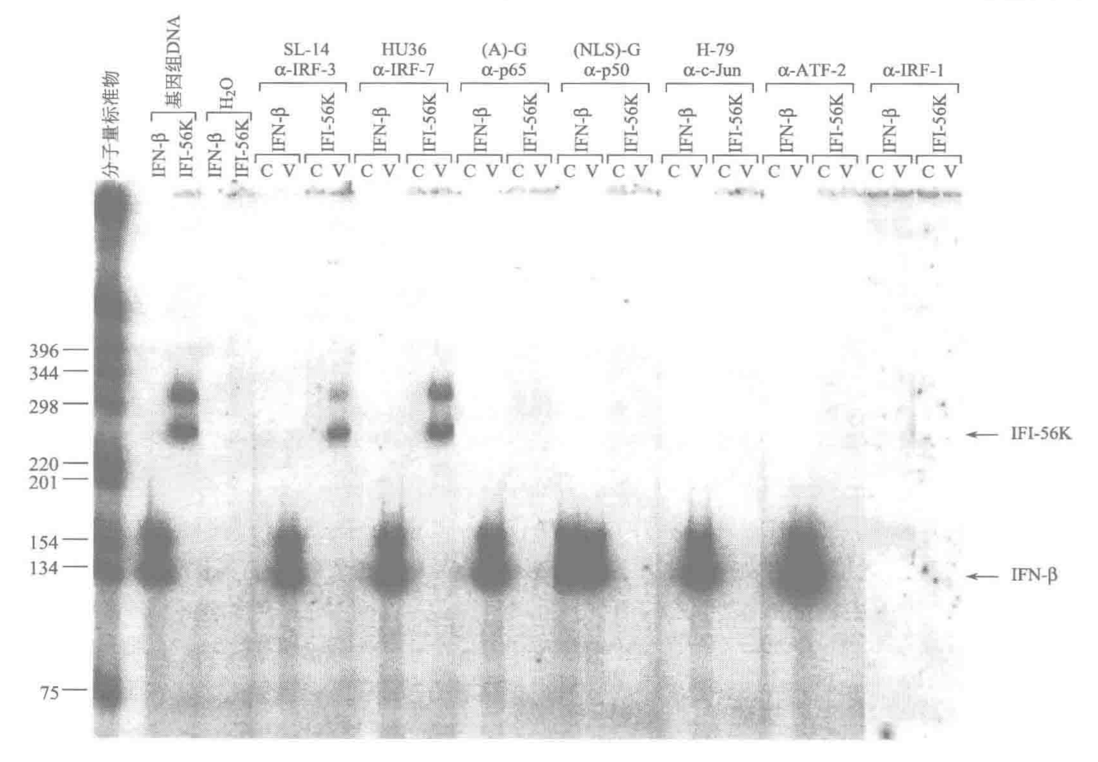


图 8.2 特定转录激活因子蛋白与病毒可诱导启动子的体内结合。HeLa 细胞模拟感染（C）或 SV（V）感染 6h，并用甲醛处理使蛋白质与结合的 DNA 交联。纯化交联物，并按之前描述的，用指出的抗体免疫沉淀（Orlando et al. 1997）。解交联之后，用 IFI-56K（−237/−219 和 +12/+31）和 IFN-β（−129/−117 和 −1/−19）的启动子引物，通过 PCR 扩增 DNA（20 个 1min 的循环，每个循环的温度为 94°C、55℃ 和 72℃）。显示在左边的是分子质量标准物和使用基因组 DNA 及水的对照扩增（经许可转载，转载自 Wathelet et al. 1998，© Cell Press）。

(p65，p50，c-Jun 和 ATF-2) 抗体的免疫沉淀确证了它们还与 IFN-β 启动子而不是对照启动子 (IFI-56K) 相互作用 (图 8.2)。这些结果为 IRF-3 和 IRF-7 的确是 IFN-β 增强子的相关激活因子提供了强大的证据。

ChIP 分析的主要优点是，它是目前可获得的用于直接“显示”特定蛋白质与内源基因控制区之间体内相互作用的唯一方法。如果在控制良好的实验中获得了有说服力的数据，研究者就可以相当自信地推断：在活细胞中特定的蛋白质结合在目的 DNA 元件附近。然而，ChIP 分析有一些局限性。第一，相对于阴性对照，ChIP 分析经常产生低信号，导致结果不完全可信。第二，由于该分析的分辨率有限，难以识别蛋白质精确的结合位点；如果 DNA 区域中包含多重潜在结合位点或完全缺乏识别位点，这个局限性尤其相关。第三，或许是最重要的，ChIP 不能证明蛋白质在其结合的控制区执行重要的功能。与这个局限性相关的是，阳性 ChIP 信号不能揭示该蛋白质是否在细胞群中与 95%的等位基因结合，或仅与 5%的等位基因发生结合；而在功能性相互作用中，一般预计蛋白质与 95%的等位基因发生结合。事实上，很多研究表明转录因子仅在与它们相互作用的一部分位点上起作用 (见 Walter et al. 1994；Li and Johnston 2001；Martone et al. 2003；Phuc Le et al. 2005；Beima et al. 2006；Hollenhorst et al. 2007；Dong et al. 2008)。一项在酿酒酵母 (*S. cerevisiae*) 中进行的研究中，Gal4 蛋白结合在含半乳糖的环境中生长的细胞中不发挥作用的基因附近，并且删除这些 Gal4 结合位点没有明显地影响细胞的生长 (Li and Johnston 2001)。另一项在哺乳动物细胞中进行的研究中，核因子-κB (NF-κB) 与许多似乎不受这个因子调控的基因结合 (Martone et al. 2003)。有趣的是，更近期的一项研究检查了含有 NF-κB 敲入突变的小鼠的表型，其中 NF-κB 敲入突变可扰乱 p300ICBP (CREB 结合蛋白) 共激活因子结合所需的磷酸化位点。这些实验揭示了几种发育上的缺陷，这些缺陷通常是由发育调控因子的错误调控引起的，而这些发育调控因子正常情况下似乎并不是 NF-κB 的靶基因 (Dong et al. 2008)。进一步的分析提示这些基因被错误调控，因为 NF-κB 偶尔结合到接近这些基因的基因组位点上，并且当 NF-κB 不能通过磷酸化正确地从一个抑制因子转变成激活因子时，NF-κB 促进这些基因的稳定沉默 (Dong et al. 2008)。

重要的是要知道 ChIP 分析中的阳性结果仅为一个因子在相互作用位点上执行重要功能的假设提供适度的支持。因此，必须探索进一步测验这个假设的其他策略 (如下所述)。在第 9 章中更详细地描述了 ChIP 分析，还讨论了重要的对照和这种分析的其他局限性。

通过基因破坏或 RNA 干扰的功能缺失研究

将转录因子与靶基因联系起来的最有效的策略之一是，运用同源重组或 RNAi 废除或敲低细胞系或动物中该因子的表达。DNA 结合蛋白的功能缺失可能导致候选靶基因表达下降，从而为蛋白质直接调控该基因提供证据。几十年来，已经具有破坏模式物种 (如大肠杆菌和酿酒酵母) 中特定基因的经典的和分子遗传学的策略。然而，直到 20 世纪 80 年代的中后期才发展出用于靶向破坏小鼠体内特定基因的同源重组技术 (Folger et al. 1984；Smithies et al. 1985；Thomas and Capecchi 1987；Capecchi 1989，2005)。

在转化的细胞系中，通过同源重组破坏哺乳动物基因也偶有成功（如见 Smithies et al. 1985）。然而，在转化的细胞系中通过同源重组破坏基因的效率一般比在小鼠胚胎干细胞（ES）中低。因此，通过同源重组的基因破坏，一般是通过破坏 ES 细胞中的一个等位基因来完成，接着产生杂合子突变小鼠，这些小鼠可被培育至纯合子。对于基因破坏实验而言，实验小鼠也优于转化的细胞系，因为可以进行突变表型的更加有意义的分析。

对于小鼠中基因的传统破坏而言，制备的打靶构体包含一抗药基因盒（一般是 *neo*^{r}），其两侧为 DNA 长片段（同源臂），这些 DNA 片段与待改变区域两侧的基因组序列同源（见第 4 章）（Thomas and Capecchi 1987；Capecchi 1989，2005）。将打靶构体转染进小鼠胚胎干细胞，然后选择抗药性，并筛选相对于非同源整合而言已经通过同源重组成功地整合打靶构体的克隆。某些研究者在打靶构体的同源臂外部包含一个单纯疱疹病毒胸苷激酶（HSV-TK）基因，该基因允许针对非同源整合的筛选，因为 HSV-TK 盒只在非同源整合时融入 ES 细胞基因组，当细胞在丙氧鸟苷中生长时导致细胞死亡（Capecchi 2005）。

在鉴定了通过同源重组整合了打靶构体的 ES 细胞克隆后，扩增克隆并注入胚囊腔，然后将其植入子宫，培育嵌合体直到检测到突变体等位基因的种系传递。大多数研究者使用同源重组生成等位基因，该等位基因中目的基因的关键区域的两侧为 *loxP* 位点，从而通过 Cre 重组酶的选择性表达，实现特定细胞类型中的基因条件性破坏（Gu et al. 1994）。关于同源重组策略的其他细节已超出了本书的范围，但可以在上面引用的文献中得到关于该策略的详细资料。

通过同源重组的基因破坏仍然是在哺乳动物细胞中进行功能缺失研究的首选策略，因为这一策略会导致蛋白质表达的彻底而特异的缺失。然而，由于小鼠中的基因破坏需要相当多的时间和费用，研究者一直在寻找更快更容易的供选方案。诸如反义 RNA（Branch 1998；Stein 1998）的策略取得了一些成功，然而，用这一方案和其他的供选方案进行成功的功能缺失实验的可能性通常很低，直到发现 RNAi。

Fire 等（1998）发现了双链 RNA 促进同源 mRNA 降解的能力（图 8.3）。在许多有机体中长的双链 RNA 促进 mRNA 的降解，包括秀丽隐杆线虫（*Caenorhabditis elegans*）和黑腹果蝇（*Drosophila melanogaster*）。然而，在哺乳动物细胞中，有必要使用短的双链 RNA 分子产物，这些 RNA 分子大小与 RNAi 加工机器生成的产物（通常 19～26 个核苷酸）类似，以避免激发针对较长分子的干扰素应答（Elbashir et al. 2001a，b）。在这个关键的发现不久之后，许多研究者通过引入或表达短干扰 RNA（siRNA）或短发夹 RNA（shRNA）成功地敲低了哺乳动物的基因（Barton and Medzhitov 2002；Brummelkamp et al. 2002；Miyagishi and Taira 2002；Paddison et al. 2002；Paul et al. 2002；Sen and Blau 2006）。现有的方法范围广泛，从合成 siRNA 的直接转染，到使用质粒、反转录病毒、慢病毒和腺病毒载体引入在 RNA 聚合酶 III 或 RNA 聚合酶 II 启动子控制下的编码 shRNA 或 microRNA 前体的基因。通过 RNAi 敲低基因表达的具体策略超出了本书的范围，但它们已经是好几百篇文章的论题（综述，见 Elbashir et al. 2002；Amarzguioui et al. 2006；Cullen 2006；Pei and Tuschl 2006；Sen

and Blau 2006；Snøve and Rossi 2006；Wiznerowicz et al. 2006；Kim and Rossi 2007；Svoboda 2007；Paddison 2008)。

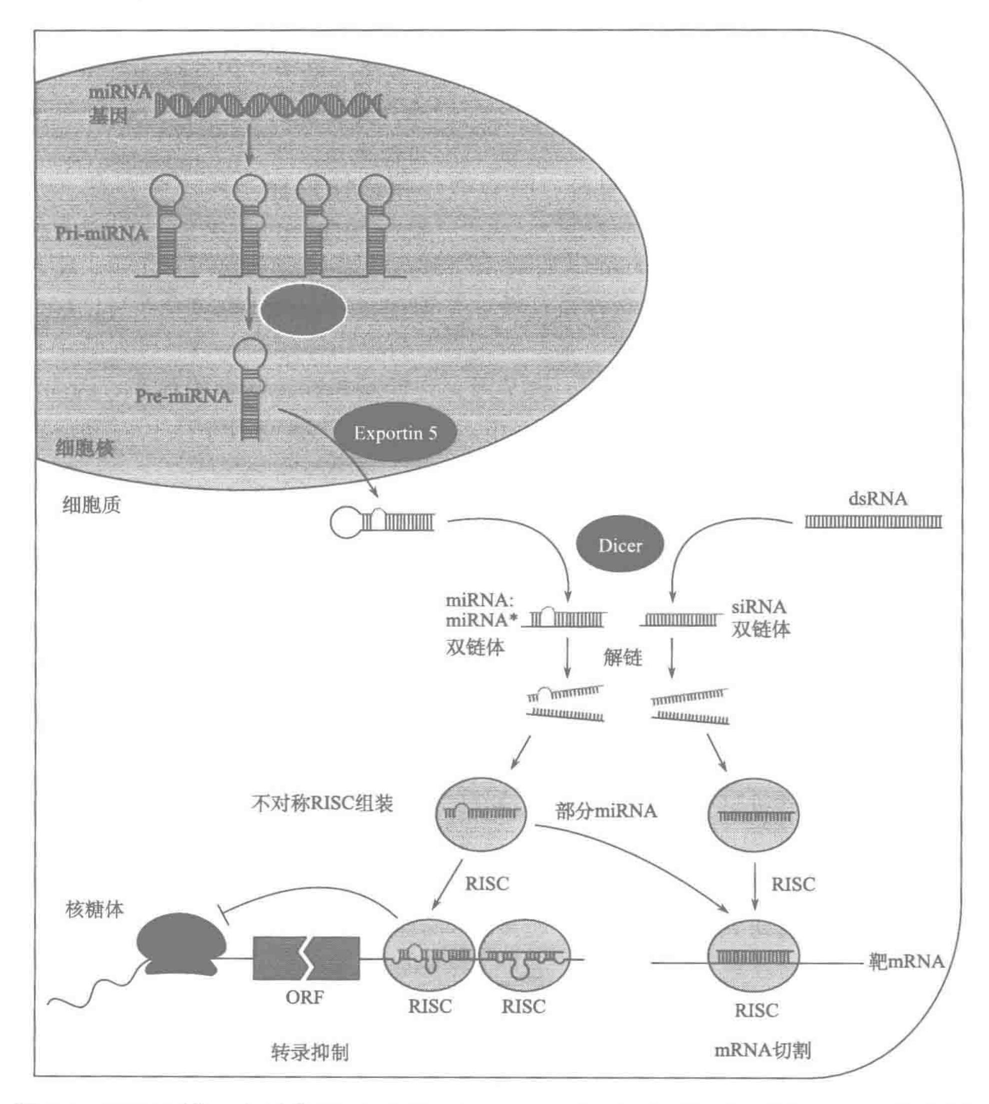


图 8.3　RNA 干扰、小干扰 RNA（siRNA）和 microRNA（miRNA）的加工及行为机制。ORE，开放可读框；RISC，RNA 诱导沉默复合物（经 Macmillan 出版公司许可，重绘自 He and Hannon 2004）。

已证明通过同源重组的 RNAi 和基因破坏对于评估转录因子对特定生物过程及潜在靶基因表达的重要性非常宝贵。然而，所有功能缺失研究的主要局限性是它难以确定哪些基因被转录因子直接调控，而哪些基因是间接靶基因。换句话说，转录因子可能与该基因的控制区直接结合，或者它能通过调控其他基因的表达间接调控该基因的表达，该基因再调节目的基因的表达。即使当将功能缺失结果与表明该因子在体内与该基因控制

区结合的阳性 ChIP 信号相结合，谨慎地解释这些数据仍然是重要的，这些组合的结果未提供这种因子直接调控靶基因的确切证据。

当通过 RNAi 而不是基因破坏进行功能缺失研究时，必须特别谨慎，因为有些证据充分的实验假象就与 RNAi 实验有关（Lassus et al. 2002；Cullen 2006；Pei and Tuschl 2006；Snøve and Rossi 2006；Svoboda 2007；Anderson et al. 2008）。特别是，siRNA 和 shRNA 经常与脱靶效应有关，原因是它们能靶向其他 mRNA 的降解，这些 mRNA 虽然不完全同源，但也有足够的同源性使这两种 RNA 与其结合。为了帮助排除一个敲低表型有可能是脱靶效应的结果，通常重要的是要表明针对同一 mRNA 的两种不同的 siRNA 产生相似的表型。此外，用一种不同的实验方法可以确证该敲低表型，如显性失活方法（见下文），或当该基因的表达已被敲低时，通过异位表达抗 siRNA 型的该基因来营救这种表型。在上面引用的论文中可找到设计 RNAi 实验和确证其结果有效性的详细方案。

许多功能缺失实验的另一个主要的局限性是，一种转录因子的破坏或敲低会对候选靶基因的表达没有影响，原因是当目的因子被消除时，另一种相关转录因子的冗余或一个相关因子的补偿性上调。这种冗余对基因破坏结果解释的影响有很好的典型，如与 T 细胞受体 α（TCR-α）增强子调控有关的淋巴增强子结合因子 1（LEF-1）和 T 细胞因子 1（TCF-1）蛋白的研究（Clevers and Grosschedl 1996；Okamura et al. 1998）。LEF-1 和 TCF-1 是高迁移率族（HMG）B 蛋白家族蛋白的高度同源的成员。两种蛋白质都能与 TCR-α 增强子中的关键元件结合，并且与结合邻近元件的 DNA 结合蛋白协作，可反式激活该增强子，同时 LEF-1 的反式激活大约比 TCF-1 的反式激活强 10 倍。然而，无论破坏 $LEF1^{-/-}$ 基因还是破坏 $TCF\text{-}1^{-/-}$ 基因，对 TCR-α 基因的转录都几乎没有影响（van Genderen et al. 1994；Verbeek et al. 1995）。如果知道只存在这种蛋白质之一，诱人的是推测这种蛋白质不是 TCR-α 增强子的相关激活因子。然而，因为两种蛋白质都被发现，通过产生两种基因都被破坏的小鼠来检查冗余的可能性。$LEF\text{-}1^{-/-}/TCF\text{-}1^{-/-}$ 小鼠显示出胸腺细胞发育中的严重缺陷，并且 TCR-α 基因转录不足，这些结果强烈暗示这些蛋白质至少部分冗余（Okamura et al. 1998）。因此，虽然基因破坏和 RNAi 实验能提供关于转录因子与候选靶基因之间关系的非常宝贵的信息，但是严格证明一种因子是靶基因的直接调控因子仍需要其他的实验证据。

体外蛋白质-DNA 复合物的丰度

虽然 ChIP 和功能缺失研究可以支持转录因子与靶基因相关 DNA 元件之间的相互作用是功能相关的假设，但是该假设的其他测验仍然是必需的。这些测验在难度水平及其结果支持该假设的程度上会有很大的变化。最简单的测验之一是评价一个 EMSA 复合物的丰度，该复合物是在混合来自合适细胞的核提取物与包含目的 DNA 元件的放射性探针时形成的。事实上，在体外结合分析中检测到一个显著的蛋白质-DNA 复合物通常形成一个初步假设的基础，该假设认为给定的蛋白质负责控制元件的功能。然而，这种初步的假设需要更多的支持。

第 6 章中论述的白细胞介素-12b（*Il12b*）启动子分析提供了一个实例，说明通过检测一个丰富的 EMSA 复合物仅能提供有限的知识。为了确定哪种蛋白质负责在－88和－99 之间的 *Il12b* 启动子元件的功能，利用来自未诱导和诱导巨噬细胞的提取物进行 EMSA 实验（Plevy et al. 1997；Bradley et al. 2003），结果检测到几种不同的蛋白质-DNA 复合物。将特定的抗体加入结合反应中，结果在丰度最高的复合物中鉴定出 CCAAT 增强子结合蛋白质（C/EBP-β），而丰度较低的复合物包含 C/EBP-δ 和 C/EBP-α，因此猜测 C/EBP-β 负责－88/－99 元件的功能。然而，这些数据不能排除一种可能性，即 C/EBP-δ、C/EBP-α 或其他未产生可检测的 EMSA 复合物的蛋白质也是功能相关的。C/EBP-β 产生丰度最高的复合物有许多原因（见表 8.1），然而，一种不同的蛋白质也可能在体内是功能相关的。例如，通过以最佳的物理或功能方式与其他结合该启动子的蛋白质相互作用。此外，在细胞提取物中丰度较低（应因此难以检测）的蛋白质可能在体内以更高的亲和力与该元件结合，或者在所用的体外 EMSA 环境中检测不到的蛋白质可能在体内优先与控制元件结合。对于 *Il12b* 启动子分析来说，要确定哪个 C/EBP 家族成员是功能相关的尤其困难。虽然丰度最高的 EMSA 复合物包含 C/EBP-β，而 C/EBP-δ 结合相同的序列，并且和 C/EBP-β 一样，其丰度也随巨噬细胞的活化而增加。

尽管多蛋白质家族的存在显著地增加了确定特定蛋白质-DNA 相互作用功能相关性的挑战性，但是要鉴定来自不同家族中的相关蛋白质也会很难。未成熟淋巴细胞中表达的末端转移酶（Dntt）基因的分析为这个问题提供了例子。DNase I 足迹法从未成熟淋巴细胞核提取物中鉴定出了一种与 Dntt 启动子中称为 D′ 元件的重要控制元件高效结合的蛋白质（Lo et al. 1991）。在纯化和克隆后发现这种蛋白质是伊卡洛斯（Ikaros），表明 Ikaros 或 Ikaros 家族成员是通过 D′元件激活 Dntt 转录的相关激活因子（Georgopoulos et al. 1992，Hahm et al. 1994）。然而，随后的实验表明转录因子 Ets 家族成员同样能以很大的亲和力与 D′元件结合，即使这些蛋白质在用核粗提物的 EMSA 或足迹实验中不易检测（Ernst et al. 1993）。用本章描述的许多测验方法，可以有力地支持 Ets-D′相互作用与 Dntt 活化的相关性，却不支持 Ikaros-D′相互作用的相关性（Ernst et al. 1996）。尽管几项证据表明 Ikaros 不是通过这个元件转录的相关激活因子，但是仍不清楚为什么 Ikaros 是表达 Dntt 细胞的核提取物中占优势的 D′结合蛋白。

DNA 结合蛋白和靶基因的相对表达模式

另一种简单的测验蛋白质-DNA 相互作用具有功能相关性假设的方法是，比较结合的蛋白质与其假定靶基因的表达模式。通过首先定量来自假定靶基因的稳态 mRNA 水平或新生转录本（见第 2 章），可以以转化的细胞系和原代细胞进行这种验证。之后，通过利用免疫印迹、免疫沉淀、免疫荧光、流式细胞仪或生化分析测量稳态蛋白质的水平，可以监测建议的调控蛋白的表达模式。

可以获得的用于监测每种 DNA 结合蛋白和靶基因的方法呈多样化。然而，首选的比较将涉及一些分析方法，这些分析方法对于证明活性 DNA 结合蛋白和假定靶基因之间的关系最有意义；这些分析方法包括定量蛋白质 DNA 结合活性（或转录激活的能

力）的生化分析、核连缀转录分析，以及监测靶基因新生转录的实时 RT-PCR 分析。DNA 结合活性可通过 EMSA 或 DNase I 足迹法监测。对于为了其刺激转录能力需要特定转录后修饰的转录因子，用针对修饰后蛋白质的抗体进行 Western 印迹也许有用，可以进一步获知哪些细胞含有活性蛋白质。作为核连缀分析的替代，通过用跨越外显子-内含子连接点的引物进行 RT-PCR，可以监测新生前体（未拼接）的转录本（见第 2 章；Lipston and Baserga 1989；Reddy et al. 1994；Bradley et al. 2003）。监测新生转录本而不检测成熟 mRNA 的原因是新生转录本的水平通常更加精确地反映转录水平，因为据认为稳定性的改变强烈地影响 mRNA 的丰度。当然，在许多情况下，通过 RT-PCR 分析 mRNA 的丰度应该足以监测靶基因的表达。通过仔细比较 DNA 结合蛋白和靶基因的表达模式所获得的信息，能支持或帮助排除假设。如果目的 DNA 结合蛋白质在表达靶基因的所有细胞类型中都存在，这个结果将支持该蛋白质是靶基因相关激活因子的假设。此外，如果鉴定细胞表达靶基因而不是表达假定激活因子，则该假设会被削弱。然而，不能否定这个假设，因为某些基因在不同的细胞类型中可被不同组合的因子激活（Lauring and Schlissel 1999）。

重要的是要注意靶基因很少在所有表达相关转录激活因子的细胞类型中表达，更有可能的是，它仅在部分细胞中表达。根据基因调控的基本组合原理（见第 11 章），一种给定的 DNA 结合蛋白通过与各自具有其自身独特表达模式的几种其他转录因子的配合作用，以变化的表达模式促成许多基因的转录。另外，正如上面所提到的，许多转录因子都受转录后机制的调控，使它们在不表达相关靶基因［如 NF-κB、激活 T 细胞核因子（NFAT）和核激素受体］的细胞中以非活性状态存在（或存在于不恰当的亚细胞区室中）。虽然通过使用监测活性转录因子丰度的分析方法可以部分解决后一个问题（见上），但是几乎从来观察不到转录因子丰度与靶基因之间完美的相关性。另一个局限性是激活因子和靶基因之间的相关表达未提供二者之间的直接功能联系，这种相关可能是偶然的，或者通过调控基因的直接调控因子之一，该激活因子可能间接调控靶基因。

对于在细胞系或组织中可诱导的基因，基本表达模式的分析可以通过比较转录因子和靶基因诱导的动力学来加强。可以用合适的诱导剂处理细胞，并比较核提取物中 DNA 结合活性的诱导时间进程和靶基因（或整合进染色体的报告基因）的诱导时间进程。如果诱导的动力学相似或所诱导的 DNA 结合活性稍早于靶基因，则该蛋白质可能的确调控这个基因。某些类别的转录因子，如核激素受体，其动力学实验可为蛋白质-DNA 相互作用具有功能相关性提供有力的证据。

解释动力学实验时需要补充说明的是，激活靶基因所需的 DNA 结合蛋白的精确浓度通常是未知的。例如，如果当 DNA 结合蛋白的浓度仅达到其最大浓度的 1%时靶基因被激活，那么该结果初步表明靶基因的诱导不依赖于该激活因子。有些情况下，为了靶基因的活化，必须出现关键 DNA 结合蛋白的高阈值浓度（Fiering et al. 1990）。如果蛋白质的阈值浓度确实很高，动力学分析就能提供信息。然而，在缺乏阈值浓度的相关信息时，难以评价试图将转录因子和靶基因诱导动力学联系起来的实验。

蛋白质结合所需核苷酸和调控元件活性所需核苷酸之间的相关性

另一种蛋白质-DNA 相互作用功能相关性的基本测验涉及控制元件功能所需要的核苷酸和假定转录因子结合所需要的核苷酸的比较。这是一个极其强大但未被充分利用的方法，该方法能提供相对强有力的证据，证明一个特定 DNA 结合蛋白，或至少这类蛋白质特定家族的一个成员负责控制元件的活性。

Dntt 启动子分析为这种策略提供了例子。正如以前所提到的，D′元件可以与 Ikaros 和 Ets 蛋白质都发生相互作用。Ikaros 是在表达 Dntt 细胞的提取物中观察到的占优势的 D′结合蛋白，表明它可能是功能性激活因子。Ikaros 和 Ets 蛋白质的表达模式不能区分哪种蛋白质是功能性激活因子，因为两个家族的蛋白质在所有被检测表达 Dntt 的细胞中都能表达。D′元件详细的突变分析证明对鉴定相关蛋白质家族是有用的 (Ernst et al. 1996)。构建一系列的单碱基或双碱基置换的 Dntt（TdT）D′元件，并在瞬时转染和稳定转染中测验它们对启动子功能的影响及其对 Ikaros 和 Ets 结合的影响。结果（表 8.2）揭示在两种转染分析中 Ikaros 结合所需核苷酸与启动子活性所需的核苷酸明显不同。两种不同突变体增强了启动子活性，却废除了 Ikaros 的结合，另一个突变体废除了启动子活性，却没有影响 Ikaros 的结合。相比之下，启动子活性所需核苷酸与 Ets 家族的各种蛋白质结合所需的核苷酸精确地匹配。这些结果支持了 Ets 家族蛋白质是 Dntt 转录的功能性激活因子的假设。

表 8.2　置换突变对启动子功能和蛋白质与 Dntt D′ 元件结合的影响

D′序列[a]											启动子功能	Ets 结合	Ikaros 结合
野生型	G	C	A	G	G	A	A	G	T	T	G	+	+	+
m83	G	$\underline{A}$	A	G	G	A	A	G	T	T	G	+/−	+/−	+
m84	G	C	$\underline{C}$	G	G	A	A	G	T	T	G	+++	++	−
m85	G	C	A	$\underline{T}$	$\underline{T}$	A	A	G	T	T	G	−	−	−
m86	G	C	A	G	G	$\underline{C}$	$\underline{C}$	G	T	T	G	−	−	−
m87	G	C	A	G	G	A	A	$\underline{T}$	T	T	G	−	−	+
m88	$\underline{A}$	C	$\underline{C}$	G	G	A	A	G	T	$\underline{A}$	G	+++	++	−

[a] 下划线为置换的核苷酸。

然而，这个策略具有很大的局限性。特别是，虽然它能解决不同蛋白质家族的相关性，但是它一般几乎不能提供同一家族不同成员间的差别。这个问题对如 Ets 家族这样的家族尤其严重，其中大量的家族成员表现出非常相似或相同的 DNA 结合特异性，并且几种家族成员在大多数细胞类型中都表达 (Hollenhorst et al. 2004，2007)。另外，用于监测突变效应的分析（如转染分析）是人工的（见第 6 章）。在上面讨论的例子中，在转染分析中 D′ 元件的活性所需要的核苷酸可能反映内源基因对 D′ 活性的需求。然而，用于转染分析的质粒高拷贝数和从其自然环境中去除控制区可能改变该分析的结果。最后一个局限

性是，仔细的突变研究更加难以评价元件的功能是否依赖于与两种或更多的蛋白质的同时结合。在这种情况下，会非常难以使活性所需要的核苷酸和结合单个蛋白质所需要的核苷酸相互关联。尽管存在这些局限性，这种策略经常提供有用的信息。

DNA 结合蛋白的过量表达对报告基因或内源基因的反式激活

过量表达或异位表达的蛋白质反式激活受目的控制区调控的报告基因或反式激活内源基因的能力，能为蛋白质-DNA 相互作用的功能相关性提供支持。然而，这些实验难以解释 DNA 结合蛋白什么时候以高于正常细胞中发现的浓度表达。在转染的细胞中报告质粒多拷贝的存在能引起类似的解释问题。

一种常见的方法开始于将编码 DNA 结合蛋白的 cDNA 插入导入培养细胞后驱动表达的载体中（图 8.4）。对于哺乳动物细胞而言，常见的表达载体包含病毒强启动子/增强子，如来自巨细胞病毒的表达载体 [如 pcDNA3.1 (Invitrogen)]。用这种表达质粒和通过目的控制区调控的报告载体共转染培养的细胞（见第 4 章），用报告基因分析监测过量表达的蛋白质对控制区活性的影响。如果过量表达导致控制区的激活，可以通过用包含结合位点突变体的报告质粒进行重复实验来评价对蛋白质结合位点的需要。在这种实验中，不应该观察到反式激活。

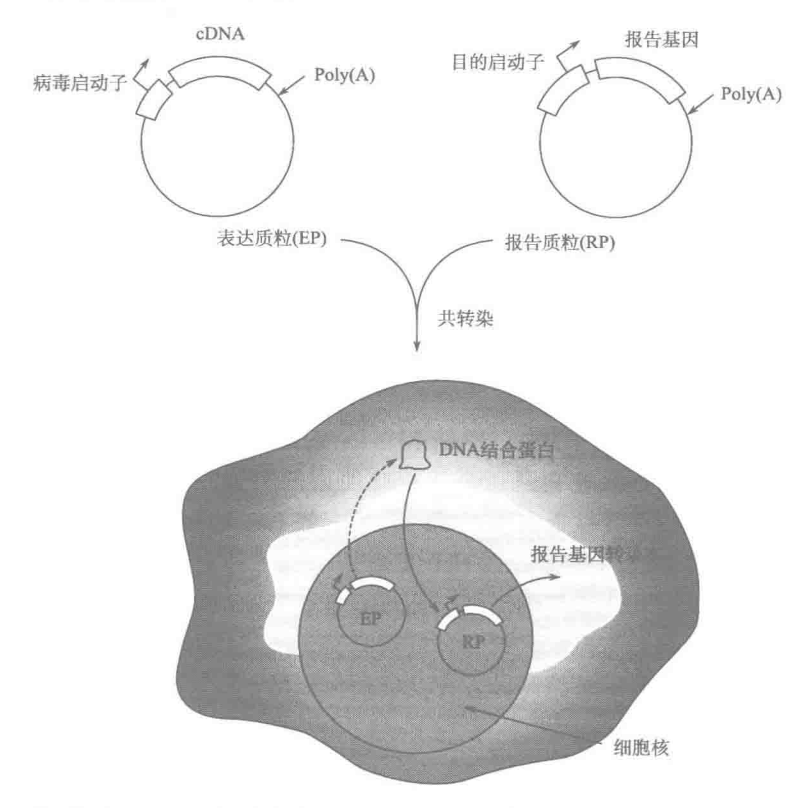


图 8.4 通过转录因子的异位过表达反式激活报告基因的表达（经允许转载自 Lodish et al. 1996，© W. H. Freeman & Company 1986、1990、1995、2000、2008）。

这类实验的阳性结果强烈暗示当 DNA 结合蛋白和控制区都过量产生时，DNA 结合蛋白能激活控制区。然而，这不能证明在生理浓度表达时蛋白质能够调控二倍体细胞中存在的两个拷贝的内源靶基因。目前的模式建议基因受多种蛋白质-DNA 和蛋白质-蛋白质相互作用的调控（见第 1 章）。通过极大地增加通常不参与调控基因的蛋白质的浓度，可能会发生异常的蛋白质-蛋白质和（或）蛋白质-DNA 相互作用，这些异常的相互作用足以激活或抑制基因。包含目的控制区的报告质粒的过量表达也会增强通常不发生的蛋白质-DNA 相互作用。

Il12b 启动子分析提供了这种方法的一个例子（Plevy et al. 1997）。为测验 C/EBP-β 是一个通过－88/－99 元件起作用的相关激活因子的假设，用 C/EBP-β 表达质粒和 *Il12b12* 启动子-报告质粒共转染未诱导的巨噬细胞。过量表达的 C/EBP-β 增强了未诱导细胞中的启动子活性，使活性达到与缺乏过量表达的诱导细胞中观察到的启动子活性类似的水平。C/EBP-β 结合位点的突变废除了反式激活。这些结果似乎暗示 C/EBP-β 是 *Il12b* 启动子的相关激活因子。然而，它们仅表明当 C/EBP-β 和 *Il12b* 启动子-报告质粒以异常高的浓度存在时启动子能被激活。这个结果没有提供重要的证据证明当以生理浓度存在时 C/EBP-β 是内源 *Il12b* 基因的相关激活因子。

当发现过量表达的转录因子反式激活内源基因时，必须考虑这些问题，然而，在这种情况下，阳性结果可能为被测验的假设提供适度水平的支持。首先，如果过量表达的 DNA 结合蛋白能反式激活正常不表达这种基因的细胞类型中的内源基因，它肯定能执行蛋白质-DNA 和蛋白质-蛋白质相互作用，这些相互作用有足够的特异性和亲和力克服染色质障碍等。其次，因为在二倍体细胞中仅存在有两个拷贝的内源基因，由不是生理相关的蛋白质引起强烈激活似乎不太可能；要取得来自内源基因的显著水平的转录，可能需要该蛋白质与其他蛋白质的高度特异性的相互作用。

一个过量表达问题的潜在的解决办法是，异位表达与正常细胞中发现的浓度相当的蛋白质。通过稳定地表达蛋白质，这个方法可能行得通。之后，分离单个的细胞克隆并且测验蛋白质的表达水平。在特定水平上表达该蛋白质的克隆可以用于靶基因转录的进一步分析，该特定表达水平与自然表达这种蛋白质的细胞系中发现的该蛋白质表达水平类似。虽然这可能提供有用的信息，但是单个 DNA 结合蛋白在正常浓度上的异位表达不可能足以激活靶基因，除非它真正是从正在使用的细胞中缺失的基因的唯一调控因子。如果需要其他组织特异性蛋白质，基因将不会被有效地激活。

能解决蛋白质过量表达顾虑的另一个实验是，当每种蛋白质过量表达到相似程度时，比较转录因子家族几个成员的活性。如果只有一个家族成员反式激活控制区，那么它必定以异常高的亲和力与控制区结合，或者它必定能实现和其他蛋白质的特异性相互作用，而这些其他蛋白质则是控制区发挥功能所需要的。

与相邻控制元件结合的蛋白质间的协作结合和协同效应

蛋白质与和控制区相互作用的其他蛋白质的协作结合的选择能力，有时可为蛋白质-DNA 相互作用的功能相关性提供支持。蛋白质与其他蛋白质功能性协同作用的选择能力，也可为蛋白质-DNA 相互作用的功能相关性提供支持。

免疫球蛋白（Ig）κ3′增强子上特定的蛋白质相互作用提供了一个协作结合的例子，这种协作结合可作为功能性蛋白质-DNA 相互作用的证据（Pongubala et al. 1992；Eisenbeis et al. 1995）。该增强子内的两个关键元件彼此紧密相邻，并且似乎协同作用。用核提取物和跨越两个元件的放射性标记探针进行 EMSA 实验，结果产生一复合物，该复合物包含结合到两个位点上的蛋白质（图 8.5，泳道 1）（Pongubala et al. 1992）。两种蛋白质对 DNA 的结合是高度协作的，任何一个位点的突变都会大大降低蛋白质与探针的结合（图 8.5，泳道 4～8）。在复合物中，发现一种蛋白质是 Ets 蛋白，命名为 PU.1；另一种蛋白质被鉴定为 IRF-4，被称为 Pip（Eisenbeis et al. 1995）。这种包含与两个功能性重要位点协作结合的蛋白质的稳定复合物的检测，为这些蛋白质（并且不是 Ets 和 IRF 家族的其他成员）是 Igκ 增强子的相关调控因子提供了有力的证据。正如前面所讨论的，用分离的控制元件检测到 EMSA 复合物只能为蛋白质具有功能相关性提供相对薄弱的证据，因为蛋白质-DNA 相互作用的体外检测标准和相关的体内相互作用的标准大不相同。然而，如果用粗提取物观察到 EMSA 复合物，并且该 EMSA 复合物包含与体内功能性协同的两个位点协作结合的蛋白质，则该复合物中的蛋白质具有功能相关性的可能性就大大增加。

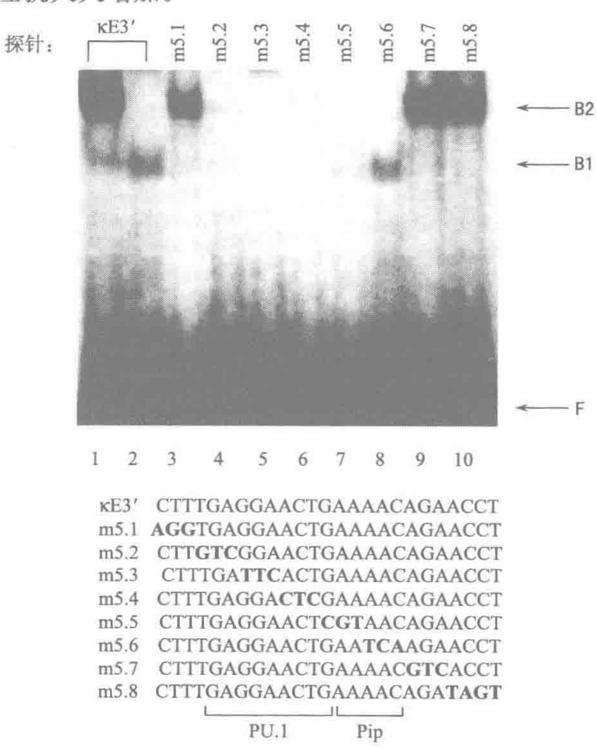


图 8.5　Igκ 3′增强子上 PU.1 和 IRF-4（Pip）的高度协同性结合。用包含图底部所示野生型或突变型增强子序列的放射性标记寡聚核苷酸进行 EMSA 实验。泳道 1 和 3～10 是用来自 S194 细胞系的核提取物进行的结合反应。泳道 2 中的结合反应包含体外翻译的 PU.1。条带 B2：PU.1 和 IRF-4 协同结合于该探针；条带 B1：单独的 PU.1；条带 F：游离探针。图的底部描绘了 PU.1 和 IRF-4（Pip）结合位点的位置（经允许改编自 Pongubala et al. 1992，©美国微生物学学会）。

这种分析的主要局限性是仅观察到一小部分蛋白质-DNA 相互作用具有强的协作结合和功能的协同作用。如果在体外没有观察到与控制区的协同结合，这种策略将毫无用处。PU. 1/IRF-4 的例子特别强大，因为在使用核粗提取物的 EMSA 实验中容易检测到协作结合的蛋白质。如果仅当 EMSA 实验中加入两种重组蛋白质时才观察到协作结合，这种结果将为该假设提供很薄弱的支持。在这种情况下，通过比较不同的 Ets 和 IRF 家族成员，表明协作结合的选择性是必需的。

IFN-β 增强体提供了协作结合和协同激活的第二个杰出的实例，该实例突出了这种方法的优点和局限性（Thanos and Maniatis 1995a，b；Kim and Maniatis，1997；Wathelet et al. 1998；Panne et al. 2007）。早期的研究发现，IRF-1 与一个功能性重要的增强子元件相互作用，并且与其他转录因子，包括 ATF-2/c-jun 和 NF-κB，协作结合于该增强子（Thanos and Maniatis 1995b；Kim and Maniatis 1997）。IRF-1 也能与那些因子协同反式激活该增强子。这些数据表明，IRF-1 是 IFN-β 转录的相关激活因子。然而，后来由 Wathelet 等（1998）用 ChIP 所做的一项研究（上面讨论的）却极力反对该假设。该研究中，在 IFN-β 转录诱导之后，鉴定了一个与 IRF 识别元件相互作用的蛋白质复合物。然而，可诱导的复合物不包含 IRF-1，而是另外两个 IRF 家族成员，即 IRF-3 和 IRF-7。另外的实验也确证了有助于体内 IFN-β 转录激活的是 IRF-3 和 IRF-7，而不是 IRF-1。特别是，该增强子中单碱基对的替换废除了 IRF-1 的结合，但是对IRF-3/IRF-7 的结合或增强子活性没有影响。因此，尽管最初的证据表明 IRF-1 能与其他相关因子协作结合该增强子，并与那些因子合作，协同性地激活该增强子，但是它似乎和增强子活性或 IFN-β 转录不相关。

最初的 IRF-1 研究存在误导性，主要的原因是使用重组蛋白质显示协作结合和协同激活。这与 PU. 1/IRF-4 的研究截然不同，在 PU. 1/IRF-4 的研究中，协作相互作用首先是在表达 Ig κ 基因的细胞的核粗提物中观察到的；不需要重组蛋白质和过量表达的蛋白质。当在诱导细胞（缺乏蛋白质过量表达）的核粗提物中以 IFN-β 增强子最终观察到协作相互作用时，发现复合物含有 IRF-3 和 IRF-7，而不是 IRF-1。

基因组足迹模式和体外足迹模式的比较

假定一种因子与一个特定控制区相互作用，则 ChIP 分析对于评估该因子在体内结合该控制区的可能性很有用。然而，ChIP 分析并不提供关于一个细胞群中与目的因子结合的等位基因份额的定量信息。如果认为两种不同的因子与相同的元件相互作用，并且两种因子都在 ChIP 实验中产生阳性结果，这可能会出现问题。在这种情形下，有可能这些因子中的其中一种可能偶然结合，而且没有功能相关性，同时仅占非常小的等位基因份额。此外，一种因子可能会产生阳性的 ChIP 信号，因为它在控制区附近相对非特异性地相互作用，而不是特异性地识别目的 DNA 元件。

一种洞悉在细胞群中以最高的效率与一个位点结合的 DNA 结合蛋白质身份的方法是比较两种基因组足迹模式，其中一种基因组足迹模式是用活跃表达靶基因的细胞观察

到的，另一种基因组足迹模式是在体外用一种已知的 DNA 结合蛋白观察到的。这种比较可用第 9 章中描述的任何一种基因组足迹法进行，包括 DNase I 足迹或硫酸二甲酯(DMS) 保护法。如果在体外观察到的足迹模式和在体内观察到的足迹模式相匹配，在体外分析的蛋白质可能确实是当它有活性时与内源基因高效结合的蛋白质，就暗示该蛋白质就是相关性激活因子。如果用 DNase I 足迹作比较，那么在体外和基因组足迹程序中一般有必要使用连接介导 PCR（LM-PCR）技术（见第 9 章），因为难以对通过 LM-PCR 获得的足迹结果和使用放射性标记探针获得的结果进行比较。相比之下，如果用 DMS 足迹，在体内通过 LM-PCR 获得的结果可以直接与体外利用放射性标记探针获得的结果比较。

在淋巴细胞特异性 Rag2 启动子分析过程中就使用了这种策略（Lauring and Schlissel 1999）。体外 EMSA 研究表明 B 细胞激活因子 Pax5 能与一个重要的控制元件相互作用。这些结果表明 Pax5 是 B 细胞内 Rag2 转录的重要激活因子。通过 DMS 基因组足迹观察到了该元件上的体内相互作用，但是体内足迹结果却没有显示相互作用蛋白质的身份。为了证明与该位点在体内结合的蛋白质是 Pax5，用重组蛋白在体外进行了 DMS 保护分析。结果显示体内和体外相互作用模式相同，这表明 Pax5 确实是占据 B 细胞中内源性元件的蛋白质。

这个策略的主要局限性是其结果一般不能区分一个蛋白质家族的各种成员，这是因为大多数蛋白质家族成员形成相似的足迹。此外，某些 DNA 结合蛋白质不能产生足够独特的足迹模式，使得难以确定与体内内源位点结合的蛋白质是否与正在体外测验的已知蛋白质相关。然而，如果获得有特征的清晰的基因组足迹和体外足迹，这种比较能为特定蛋白质家族在确定的控制元件的功能中的作用提供支持。

蛋白质-DNA 相互作用的相对亲和力

能支持其相关性的蛋白质-DNA 相互作用的特性之一是它的亲和力，该亲和力是相对于可在相同位点上发生的其他蛋白质-DNA 相互作用的亲和力而言的。据我们所知，还没有对转录因子的亲和力与功能相关性之间的关系进行过细致的和系统的检查，尽管在模式系统中，转录因子对其靶位点的亲和力与活化成正比（见 Mauxion et al. 1991; Lehman et al. 1995)。某些情况下（如前面所描述的 PU. 1/IRF-4 例子），单个的 DNA 结合蛋白质倾向于以低亲和力相互作用，以致混合元件的活性依然依赖于两种或更多种蛋白质的协作结合。然而，在许多情况下，相对于能结合相同位点的不相关蛋白质的亲和力，与控制元件发生功能性相互作用的蛋白质似乎有可能将以异常高的亲和力与该控制元件结合。

为了确定多蛋白家族中哪个成员最有可能和一个元件进行功能性相互作用，亲和力也许是特别有用的标准。虽然多蛋白家族的各种成员可能识别非常相似的序列，但是它们经常有细微的序列偏好。在确定哪个因子作用于特定的控制元件的过程中，这些偏好发挥重大作用。如果这种假设是正确的，仔细比较多蛋白家族的成员对一个位点的结合亲和力可能是有用的。

有数种方法可用于比较各种蛋白质对特定位点的相对亲和力。获得纯化的重组蛋白质并进行仔细的平衡解离常数的测定（见第 11 章），这种方法提供了有价值的信息，但有两个不利之处。首先，难以准备并精确定量应测验的所有重组型蛋白质，特别是当研究一个大型多蛋白家族时。更重要的是，重组蛋白质对于 DNA 序列的亲和力可与天然蛋白质的亲和力显著不同。原因之一是翻译后修饰或与其他蛋白质的相互作用可改变蛋白质-DNA 相互作用的亲和力。特别是，正发现越来越多的蛋白质拥有自我抑制结构域，除非被适当地修饰，否则这些结构域会降低蛋白质对 DNA 的亲和力（见 Jonsen et al. 1996；Lee et al. 2005；Pufall et al. 2005）。

一个能提供更多信息的实验将是比较在细胞提取物中的相对亲和力，因为在细胞提取物中蛋白质更可能以它们的天然状态存在。如果提取物中大多数目的蛋白质的浓度足够高，且可以在基本 EMSA 中检测到，就能获得它们对于一个位点的亲和力的近似的比较。一种方法是通过非特异性竞争 DNA、盐或去污剂在结合反应中的滴定，确定每种蛋白质-DNA 复合物对这些试剂的感受性。较高亲和力的相互作用对这些试剂可能更有抵抗力，特别是在比较家族的各成员时。遗憾的是，这种方法仅能比较那些浓度足够高，可在基本 EMSA 实验中检测的蛋白质。而对于那些相对低丰度或其结合特性与所使用的 EMSA 条件不相容的蛋白质，则不能测验。

一种可替代的方法是用上面讨论的 Dntt D′ 元件进行的（图 8.6）（Ernst et al. 1996）。目的是确定表达 Dntt 的细胞系的提取物中哪种 Ets 蛋白质以最高的亲和力与 D′ 元件结合（如为 Dntt 转录的功能性激活因子鉴定候选者）。在 Dntt 表达的细胞中存在几种已知的，以及或许其他新的 Ets 蛋白，但是用核提取物仅能检测到几种包含

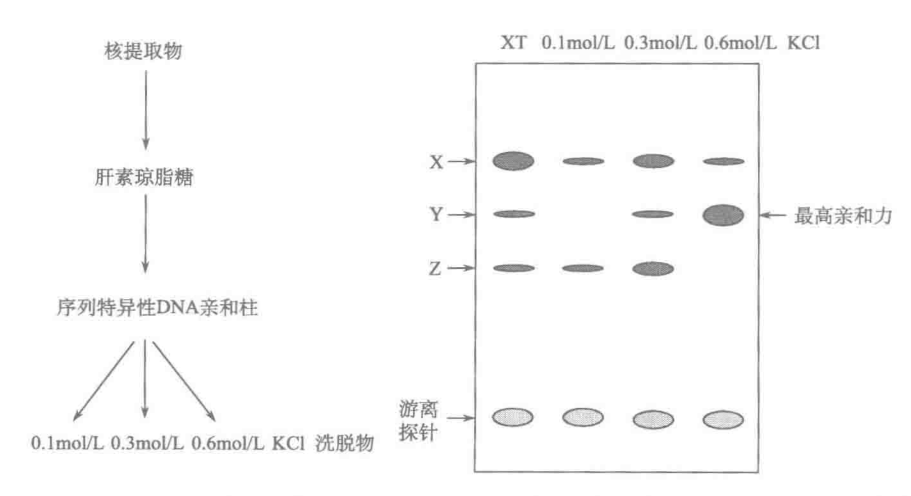


图 8.6　通过序列特异性 DNA 亲和层析，比较核粗提物中蛋白质-DNA 相互作用的亲和力。首先将来源于鼠 T 细胞系 RLm11 的核提取物通过肝素琼脂糖分级，去除核酸酶。然后用带有共价连接的包含 Dntt 启动子 D′ 元件的寡核苷酸树脂进行序列特异性 DNA 亲和层析。样品加到含有 0.1 mol/L KCl 缓冲液的柱子上。然后用 0.3 mol/L 和 0.6 mol/L KCl 洗脱蛋白质。图片右侧描绘了理想的 EMSA 结果；将提取物用于泳道 1 的结合反应时，检测到三种蛋白质-DNA 复合物，但直到 0.6 mol/L KCl 洗脱时，仅有一种复合物（复合物 Y）优先保留在柱上（泳道 4）。

Ets 蛋白的 EMSA 复合物。为了避免偏向于已知蛋白质或用粗提物通过 EMSA 检测到的蛋白质，可以使用序列特异性 DNA 亲和层析进行处理（见第 7 章）。预期亲和力最高的蛋白质能以最高的盐浓度从亲和柱中洗脱。确实，亲和层析之后，在高盐洗脱物中观察到单一高丰度 EMSA 复合物，所有其他 EMSA 复合物在低盐洗脱物中丰度更高。银染的蛋白质凝胶鉴定了负责该复合物的蛋白质（图 8.6，复合物 Y），并且肽测序揭示了该蛋白质是一种称为 Elf-1 的 Ets 蛋白质。柱分级物的免疫印迹确证了 Elf-1 比其他几种 Ets 家族成员在更高的盐浓度下被洗脱。随后的实验为 Elf-1 是通过 D′ 元件的 Dntt 转录的功能性激活因子提供了另外的支持。然而，该假设仍未证实。因此，目前还没有从本研究或其他实验室进行的研究中获知在密切相关的蛋白质的大家族中亲和力对评价功能相关性是否有效。

亲和层析法具有几个超过上面提及的其他亲和力策略的优点。首先，在粗提物中自然蛋白质更有可能保留在完整细胞中存在的翻译后修饰。其次，这种方法不偏向于丰度最高的蛋白质或以前描述过的蛋白质。实际上，这种策略应导致任何丰度适度的蛋白质的纯化，这些蛋白质与该位点以高亲和力结合。最后，如果以高亲和力结合的蛋白质是新的，那么通过质谱分析就可以鉴定该蛋白质（见第 7 章）。

亲和层析法的一个值得注意的局限性是，在提取物中可能不能保留存在于完整细胞中的翻译后修饰。此外，如果结合一个位点的不同蛋白质都不同程度地依赖于 DNA 结合结构域和 DNA 之间的静电作用，那么用亲和层析作为相对亲和力的测量会有些不准确。如果这种静电作用相当不同，来自亲和柱的盐洗脱物将不提供相对亲和力的精确评估。

显性失活突变体

根据定义，显性失活突变体是一种蛋白质变体，当这种突变体在含有野生型蛋白质的细胞中表达时，它会扰乱该蛋白质的功能。这种扰乱经常发生，因为突变体保留野生型蛋白质的某些但非全部的活性，这使它竞争野生型蛋白质的重要靶点或底物，尽管这种竞争的效率不高。DNA 结合蛋白质的显性失活突变体有时能提供蛋白质-DNA 功能重要性的有关信息。但是，这些实验所提供的信息有限，所用的显性失活突变体的类型对数据的解释有显著的影响。

进行显性失活研究的一个常见的方法是用报告质粒和表达质粒共转染细胞，其中报告质粒包含由目的控制区调控的标准报告基因，而表达质粒可表达显性失活型 DNA 结合蛋白质（图 8.7）。显性失活蛋白质经常被设计成保留结合 DNA 的能力，但不执行转录激活功能。换句话说，蛋白质包含其 DNA 结合结构域，但缺少其转录激活结构域（见第 10 章）。突变体蛋白质预期和内源性野生型蛋白质竞争，因此阻止野生型蛋白质激活报告基因的转录（图 8.7）。经常用这种显性失活蛋白质的抑制作用为野生型蛋白质功能性激活控制区融合的报告基因提供证据。

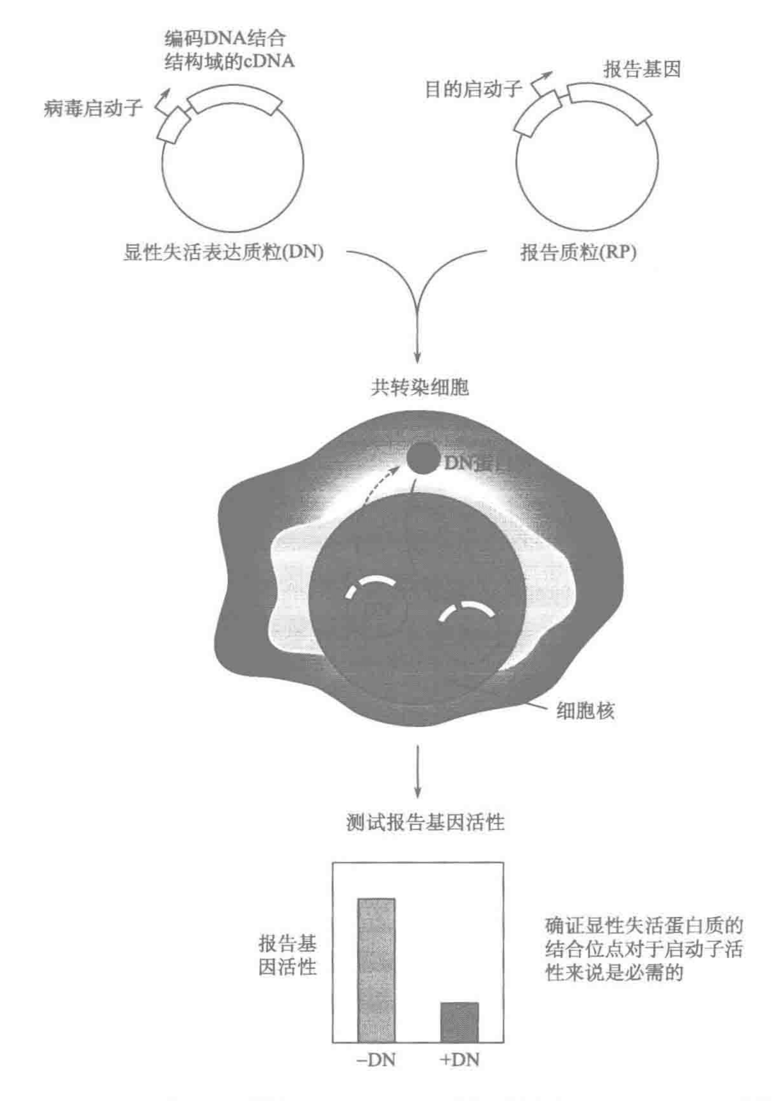


图 8.7　通过保留其 DNA 结合结构域的显性失活型转录因子抑制报告基因活性。

Plevy 等（1997）用该方法支持了 C/EBP-β 是脂多糖（LPS）活化的巨噬细胞内 *Il12b* 启动子的相关激活因子的假设。将浓度递增的称为 LIP 的显性失活型 C/EBP-β 的表达质粒（Descombes and Schibl，1991）和 *Il12b* 启动子-报告质粒共转染巨噬细胞。LIP 含有 C/EBP-β 的碱性亮氨酸拉链（bZIP）DNA 结合结构域和二聚化结构域，但缺少其转录激活结构域。之后，用脂多糖（LPS）活化细胞，利用报告基因分析监测 LIP 对 *Il12b* 启动子活性的影响，结果显示，启动子活性被 LIP 抑制。对照实验显示 LIP 的表达对包含无 C/EBP 结合位点的启动子的报告质粒没有影响。

这种类型的显性失活实验能够支持关于 DNA 结合蛋白质（如 C/EBP-β）的功能相关性的假设。然而，对这个实验的仔细考虑揭示该实验结果仅仅确证当 DNA 结合结构

域过量表达时，它能与重要的控制元件结合，并在其结合处阻断功能性激活因子的激活作用。不管功能性激活因子的身份，过量表达的显性失活突变体能简单地通过占据结合位点阻断重要的蛋白质-DNA相互作用。因此，此结果提供两条相对不太重要的信息：①显性失活蛋白质结合的控制元件对控制区的功能是重要的，该结果已经通过启动子突变分析被确立；②显性失活蛋白质中的DNA结合结构域在过量表达时能与该控制元件在体内结合。

一项替代性的实验是测验同一显性失活突变体对内源基因表达的影响。这样就有机会确证在人工转染分析中发现的重要控制元件也是重要的内源性调控者。然而，该实验存在大多数上述实验的局限性和其他两点要注意的问题。首先，显性失活蛋白有可能通过改变细胞内其他基因的表达而间接地抑制内源靶基因的转录。例如，显性失活蛋白可能抑制生存所需的基因的表达，从而导致间接抑制靶基因转录的毒性（当分析显性失活蛋白质对被转染报告质粒的影响时，也必须考虑间接影响的可能性，但这可以测试突变体报道基因加以解决或至少部分解决）。其次，分析内源靶基因在技术上比分析瞬时共转染显性失活突变体和报告质粒更难。在共转染分析中，实际上吸收报告质粒的每个细胞都会吸收显性失活表达质粒。因此，可分析整个细胞群体以考察显性失活蛋白质的影响。相比之下，为了监测对内源基因转录的影响，必须认识到只有一部分细胞将被转染并表达显性失活蛋白质。因为大多数细胞将不被转染，难以观察显性失活蛋白质对靶基因表达的影响。

然而，对于这些额外的困难，有几个相当简单的解决方法。第一，显性失活蛋白质可从能感染所有细胞的高滴度的反转录病毒中表达，可能导致对假定靶基因转录的可测量的影响。第二，显性失活蛋白质可以通过稳定转染在细胞中表达。最好的是，显性失活蛋白质应该由可诱导启动子控制，以便在选择过程中不显露这种蛋白质的任何毒性。第三，可以通过流式细胞仪或相关的免疫学方法将瞬时转染显性失活表达质粒的细胞从未转染的细胞中分离出来（见第4章）。

虽然仅保留DNA结合结构域的显性失活蛋白质用处不大，但对于解决特定问题，其他类型的显性失活蛋白质会更有用。特别是，显性失活突变体可用于研究以二聚体形式结合DNA的蛋白质，如bZIP和碱性螺旋-环-螺旋（bHLH）蛋白质。可表达属于这些类型的蛋白质的显性失活型蛋白，这些显性失活型蛋白保留二聚化结构域，但缺失DNA结合结构域。如果这些显性失活蛋白质抑制假定靶基因的转录，那么该结果可能暗示了能与显性失活蛋白质二聚化的蛋白质功能性地激活目的控制元件。依据二聚化结构域的性质，这可能牵涉一个完整的DNA结合蛋白质家族，或者这可能暗示来自家族成员特定子集的蛋白质可能发挥作用。可能有必要测试二聚化结构域与蛋白质家族的各个成员相互作用的能力。通过将突变引入使显性失活蛋白质以异常高的效率与内源蛋白形成二聚体的二聚化界面，已经为某些种类的转录因子，包括bZIP和bHLH因子，设计了异常高亲和力的显性失活蛋白质（Olive et al. 1996；Krylov et al. 1997；Rishi and Vinson 2003）。

Id1是通过与特定转录因子二聚化起作用的显性失活蛋白的经典例子（Benezra et al. 1990）。Id1和Id家族其他成员包含能使它们与bHLH激活因子形成异源二聚体的

HLH 结构域，但它们缺少 DNA 结合所需的基本区域。因此，Id 蛋白通过阻碍 bHLH 蛋白与 DNA 的结合，抑制凭借 bHLH 蛋白的转录激活。因此，Id 蛋白能用于评价 bHLH 蛋白质家族对控制元件功能的相关性。然而，因为每种 Id 蛋白质都能与多种 bHLH 蛋白形成二聚体（Langlands et al. 1997），转录激活的抑制对功能性相关激活因子的身份仅提供有限的认识。

体外转录策略

体外转录实验能支持蛋白质-DNA 相互作用的功能相关性，这种基本方法开始于控制区体外转录分析方法的发展，而这往往是非常具有挑战性的（见第 4 章、第 12 章和第 13 章）。核小体重建可帮助该分析方法的发展。为了确定目的 DNA 元件是否有助于体外活性，可以测试元件中的突变。如果这种分析依赖于 DNA 元件，那么有可能用这种分析评价特定 DNA 结合蛋白对该元件功能的相关性。作为出发点，可将抗候选蛋白的单克隆或多克隆抗体加入体外反应中，以确定这些抗体是否能阻断控制元件的功能（图 8.8A）。因为难以加入足量的抗体以中和蛋白质，所以可能需要亲和纯化的浓缩抗体。为了确定是否已加入了足量的抗体，可以监测抗体对 EMSA 复合物的影响。还应该进行对照实验，以确定抗体是否影响来自不依赖于该 DNA 结合蛋白质的无关启动子的转录。还需要其他的对照，包括监测以相似的方法制备的无关抗体的影响。如果抗体的加入不抑制转录活性，那么候选结合蛋白对控制元件的活性可能不是必需的。此外，抗体可能不结合至关重要的蛋白质表位，或其浓度不足以中和提取物中所有的蛋白质分子。

另一种可能更成功的策略是通过免疫沉淀、免疫亲和层析或序列特异性 DNA 亲和层析，从提取物中去除蛋白质（图 8.8B）（见第 7 章）。免疫沉淀和免疫亲和方法应能有效地去除抽提物中的蛋白质，除非蛋白质的关键表位被紧密结合的其他蛋白质所封闭。DNA 亲和层析导致了与目的元件结合的所有蛋白质的耗尽，从而对与元件发生功能性相互作用的蛋白质的身份几乎不提供信息。然而，通过随后加入重组或纯化形式的假定调控因子，就可以提供那种信息。如果恢复高效转录，那么该结果支持这种蛋白质是该基因相关调控因子的假设。

实际上，用于监测体外转录分析中对某种蛋白质需求的缺失实验可能乏味且困难。由于提取物的稀释或通用转录因子的失活，蛋白质缺失会非特异性地抑制体外转录。抗体或 DNA 树脂结合的蛋白质也可能与体外转录必需的其他蛋白质结合或形成聚集，从而导致这些蛋白质的共缺失。这些困难很常见，因为在缺失期间，提取物必须保持在浓缩状态，以便在随后的体外反应中起作用。

由于这类实验潜在的问题，需要设置仔细的对照，包括使用对目的启动子转录无影响的抗体的免疫缺失，以及不应受缺失影响的对照模板。此外，为了显示特定抑制只归因于目的蛋白的缺失，那么在反应体系中补充重组的或纯化的蛋白质应有可能恢复转录。当用精心设计的对照及表位易于触及的抗体进行缺失实验时，缺失实验可能相当成功。用这些体外方法获得的结果可支持蛋白质对靶基因转录的功能重要性，但不能确凿地证明这一点，而且在体外和控制元件发生功能性相互作用的蛋白质未必负责该控制元件的体内功能。

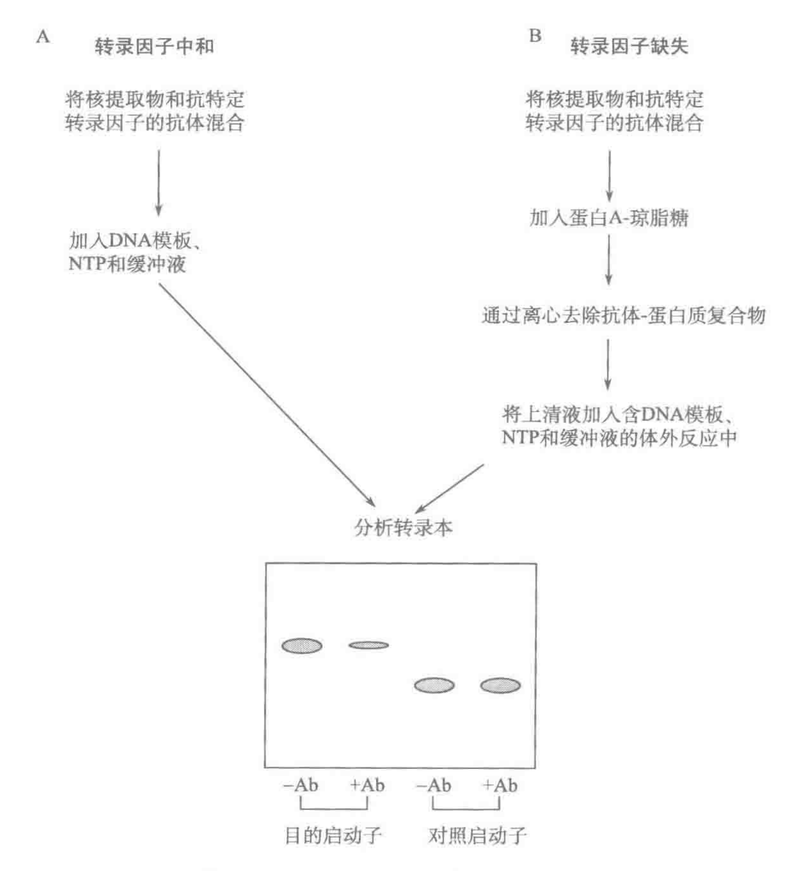


图 8.8 通过抗体对体外转录反应的抑制评价蛋白质-DNA 相互作用的相关性。

借助免疫亲和及 DNA 亲和层析的转录因子缺失被用于 IFN-β 启动子的体外转录分析（Kim and Maniatis 1997）。据认为，该启动子的体内活性需要 IRF 家族成员 ATF-2/c-Jun、NF-κB 和 HMGA（见第 11 章）。为研究启动子体外功能的特定蛋白质需求，发展了一种依赖每种因子结合位点的体外转录分析方法。之后，通过序列特异性 DNA 亲和层析缺失这三种因子。通过免疫亲和层析缺失多余的 HMGA。将每种蛋白质的重组型蛋白加回到缺失提取物中，结果会导致启动子活性的恢复。这些结果为所用的特定蛋白质能够激活启动子提供了证据。然而，应该注意到这些研究使用的是重组 IRF-1，而更近的发现强烈暗示其他 IRF 家族成员，特别是 IRF-3 和 IRF-7，实际上是 IFN-β 转录的相关激活因子（Wathelet et al. 1998；见上述）。因此，对于显示蛋白质-DNA 相互作用的真实体内相关性，独自的体外转录方法是不可靠的。

虽然本部分聚焦于粗提取物中的蛋白质中和及缺失，但体外转录也可以作为一种分析方法进行使用，以用于无偏纯化激活启动子的蛋白质（见第 12 章）。使用该策略的一个经典的例子导致了刺激 SV40 早期启动子的蛋白质的鉴定、纯化和克隆（Dynan and Tjian 1983a，b；Kadonaga et al. 1987）。在该研究中，首先发展了体外转录分析，该分析支持从 SV40 早期启动子上的精确的转录。然后，对细胞提取物进行分级分离，以鉴

定重要的启动子特异性转录因子。结果发现，所鉴定的因子，即 Sp1，能与该启动子内特定的 DNA 序列元件结合。然后纯化了 Sp1，获得了其部分肽序列，并克隆了其基因。作为体外选择性激活 SV40 启动子的蛋白质的 Sp1 的分离，为它是 SV40 启动子体内相关激活因子的假设提供了很有力的支持。

改变特异性实验

测验蛋白质-DNA 相互作用相关性的最后一种策略是改变特异性策略（图 8.9）。虽然难以设计和进行，但该方法有可能比本章讨论的其他任何策略提供更有说服力的蛋白质-DNA 相互作用相关性的证据。为了进行改变特异性实验，首先将目的蛋白的 DNA 结合结构域突变，使它识别不同的 DNA 序列。然后，将改变的蛋白质识别的新序列插入到目的控制区中代替野生型蛋白质识别的序列元件。重要的是新的识别序列要非常独特，以便不被细胞内其他的 DNA 结合蛋白以可观的亲和力识别。之后，在含有受改变的控制区调控的内源基因或报告基因的细胞内表达改变特异性的 DNA 结合蛋白，并监测其调控转录的能力。

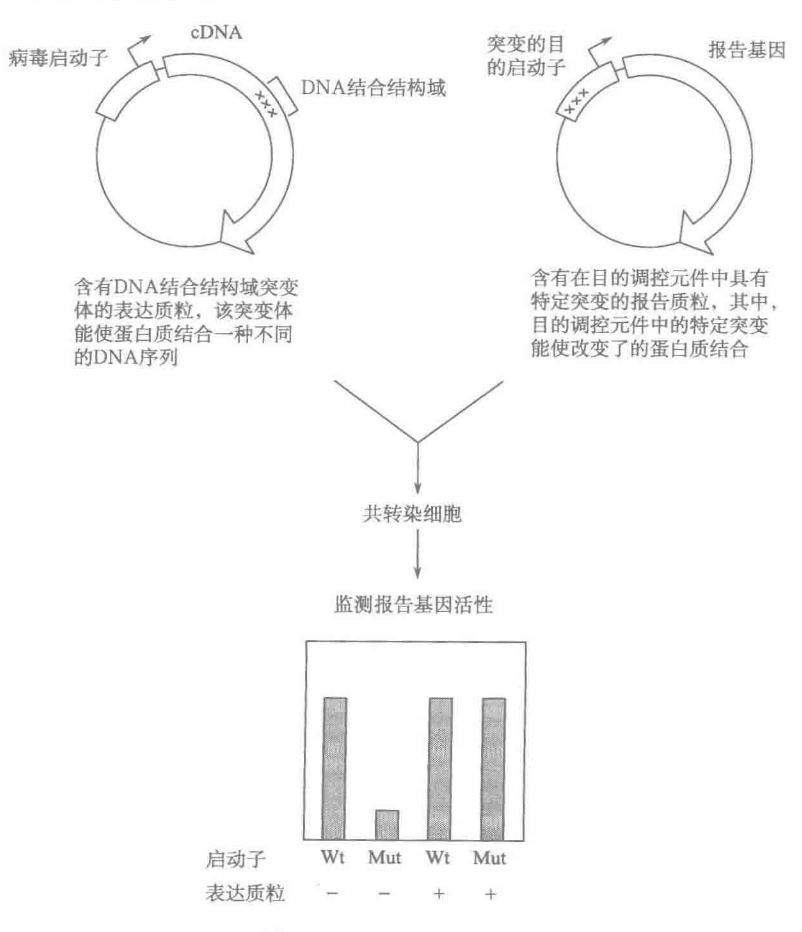


图 8.9 改变特异性策略。

Shah 等（1997）报道了改变特异性实验的一个特别好的例子，该例子揭示了这种方法的优点和缺点［Gillemans 等（1998）报道了另一个重要的例子］。该研究的目标是确定哪一个 POU 结构域蛋白质——Oct-1 或 Oct-2，能够功能性激活 Ig 重链和轻链基因启动子。这两种蛋白质都能以相似的亲和力与 Ig 启动子上八聚体元件结合，且都能在 B 淋巴细胞中表达，而基因破坏实验未提供相关的信息，因此，难以鉴定 Ig 启动子的相关激活因子。

为创制改变特异性的 Oct 蛋白，Shah 等（1997）检查了 Oct-DNA 复合物的已知晶体结构。他们聚焦于 POU 结构域中与八聚体 DNA 序列中的一个核苷酸接触的一段氨基酸序列。他们将此核苷酸突变，从而破坏与野生型 Oct 蛋白质的结合。然后，他们分离了能以高亲和力与改变的 DNA 序列结合的 Oct POU 结构域。为达到此目标，他们为 POU 结构域构建了噬菌体表达文库，其中每种噬菌体表达一种 POU 结构域，这些 POU 结构域在预测与改变的碱基对非常靠近的区域中具有随机产生的氨基酸序列。为分离表达能与改变的 DNA 序列结合的突变型 POU 结构域的噬菌体，可用包含改变序列的放射性标记寡核苷酸探测表达文库。之后，确定挑选的 POU 结构域中引起高亲和力结合的氨基酸，并将其引入表达 Oct-1 和 Oct-2 的哺乳动物表达质粒。

在 B 细胞中测试改变的 Oct-1 和 Oct-2 蛋白质反式激活一种报告质粒的能力，该报告质粒带有包含改变的八聚体 DNA 序列的 Ig 启动子及其控制下的报告基因。通过使用改变的八聚体序列，B 细胞中内源性 Oct-1 和 Oct-2 蛋白质对 Ig 启动子不赋予任何功能；只有导入细胞中的改变特异性的 Oct-1 和 Oct-2 能够与之结合。因为该实验在 B 细胞中进行，所以认为改变的蛋白质将与 B 细胞特异的 Ig 启动子和增强子活性所需的其他蛋白质发生功能性相互作用。

这些结果揭示了改变的 Oct-41 和 Oct-2 蛋白质同样能刺激 Ig 启动子活性。然而，令人惊讶的是，当报告质粒同时包含 Ig 启动子和 Ig 增强子时，Oct-1 成为更强的激活因子。这些结果与 Oct-1 是内源性 Ig 基因功能性激活因子的假设一致，因为只有 Oct-1 能进行启动子或增强子依赖性转录所需的蛋白质-蛋白质相互作用。

改变特异性策略能为 DNA 结合蛋白作用于特定的靶位点提供有说服力的证据。实际上，该方法与基因破坏实验或 RNAi 实验一样可以提供许多相同的信息，但是消除了这些实验关键的局限性。基因破坏和 RNAi 研究中，DNA 结合蛋白的缺失或敲除会导致细胞丧失生存活力。另外，极难区别结合蛋白的缺失或敲除对候选靶基因的直接和间接影响。DNA 结合蛋白也可能与相关蛋白质具有冗余性，致使其缺乏对靶基因转录没有影响。改变特异性的实验避免了这些局限性，包括：①因内源 DNA 结合蛋白仍表达，细胞保持活力；②间接影响不太可能，因为特异性改变蛋白质的改变识别位点仅被引入目的控制区中；③未观察到冗余，因为改变的 DNA 结合结构域只存在于一种蛋白质中。

尽管有这些优点，改变特异性的方法仍有三个局限性。第一，该程序一般涉及依靠瞬时转染或稳定转染的被改变 DNA 结合蛋白的过表达。正如以前所讨论的，用过表达蛋白质获得的结果是非决定性的并且难以解释。为提高改变特异性实验结果的有效性，改变蛋白质的表达浓度应与内源性野生型蛋白质的表达浓度相似。为了实现这一点，可

检查表达改变蛋白的不同稳定细胞系，或测试含有不同启动子或增强子的各种表达载体。

当在瞬时转染报告质粒环境中分析改变的控制区时，就产生第二个局限性。报告质粒的高拷贝数及其与其天然染色体环境的脱离，可能影响其对改变特异性的蛋白质作出应答的能力。理想的情况下，将创制改变控制元件的置换突变通过同源重组引入内源基因，这种严密的方法将为特定 DNA 结合蛋白作用于靶基因提供强有力证据。当然，这种实验比基本的转染实验需要更多的努力。相对于瞬时转染报告质粒的使用，稳定转染报告质粒的使用更受欢迎，但其结果仍是非决定性的。

第三个局限性是野生型 DNA 结合蛋白的转录激活功能可能依赖于被改变的特定氨基酸，这个局限性不能克服。这些氨基酸可能直接参与转录激活，或该蛋白质-DNA 相互作用能引起参与转录激活的蛋白质的另一个表面发生构象改变。在大多数改变特异性的 DNA 结合蛋白中，负责 DNA 结合的氨基酸对转录激活是不重要的，但也曾有描述 DNA 结合与激活之间构象连接的例子（Lefstin and Yamamoto 1998）。然而，尽管有这些局限性，但改变特异性实验能提供关于蛋白质-DNA 相互作用相关性的有价值的信息，并在本章所描述的策略库中占有重要的位置。

综上所述，本章我们描述了 12 种通用策略，这些策略可用于检查一个特定蛋白质是否通过与一个确定的控制元件的结合直接调控靶基因。这些策略中的某些策略对于分析某些 DNA 结合蛋白可能不可行，此时应该想到其他未描述的策略（如体内抗体微注射实验）。没有一种方法能单独决定性地确立蛋白质-DNA 相互作用的功能相关性。然而，通过尽可能多的严格测验，可以极大地巩固一个与之相关的假设。

参考文献

Amarzguioui M.，Lundberg P.，Cantin E.，Hagstrom J.，Behlke M. A.，and Rossi J. J. 2006. Rational design and in vitro and in vivo delivery of Dicer substrate siRNA. *Nat. Protoc.* 1：508-517.

Anderson E.，Boese Q.，Khvorova A.，and Karpilow J. 2008. Identifying siRNA-induced off-targets by microarray analysis. *Methods Mol. Biol.* 442：45-63.

Barton G. M. and Medzhitov R. 2002. Retroviral delivery of small interfering RNA into primary cells. *Proc. Natl. Acad. Sci.* 99：14943-14945.

Beima K. M.，Miazgowicz M. M.，Lewis M. D.，Yan P. S.，Huang T. H.，and Weinmann A. S. 2006. T-bet binding to newly identified target gene promoters is cell type-independent but results in variable context-dependent functional effects. *J. Biol. Chem.* 281：11992-12000.

Benezra R.，Davis R. L.，Lockshon D.，Turner D. L.，and Weintraub H. 1990. The protein Id：A negative regulator of helix-loop-helixbinding proteins. *Cell* 61：49-59.

Boyd K. E. and Farnham P. J. 1997. Myc versus USF：Discrimination at the cad gene is determined by core promoter elements. *Mol. Cell. Biol.* 17：2529-2537.

Boyd K. E.，Wells J.，Gutman J.，Bartley S. M.，and Farnham P. J. 1998. c-Myc target gene specificity is determined by a post-DNA-binding mechanism. *Proc. Natl. Acad. Sci.* 95：13887-13892.

Bradley M. N.，Zhou L.，and Smale S. T. 2003. C/EBPβ regulation in lipopolysaccharide-stimulated macrophages. *Mol. Cell. Biol.* 23：4841-4858.

Branch A. D. 1998. A good antisense molecule is hard to find. *Trends Biochem. Sci.* 23：45-50.

Braunatein M.，Rose A. B.，Holmes S. G.，Allis C. D.，and Broach J. R. 1993. Transcriptional silencing in

yeast is associated with reduced nucleosome acetylation. *Genes Dev*. 7: 592-604.

Brummelkamp T. R., Bernarda R., and Agami R. 2002. A system for stable expression of short interfering RNAs in mammalian cells. *Science* 296: 550-553.

Capecchi M. R. 1989. Altering the genome by homologous recombination. *Science* 244: 1288-1292.

Capecchi M. R. 2005. Gene targeting in mice: Functional analysis of the mammalian genome for the twenty-first century. *Nat. Rev. Genet*. 6: 507-512.

Clevers H. C. and Grosschedl R. 1996. Transcriptional control of lymphoid development: Lessons from gene targeting. *Immunol. Today* 17: 336-343.

Cullen B. R. 2006. Enhancing and confirming the specificity of RNAi experiments. *Nat. Methods* 3: 677-681.

Descombes P. and Schibler U. 1991. A liver-enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are translated from the same mRNA. *Cell* 67: 569-579.

Dong J., Jimi F., Zhong H., Hayden M. S., and Ghosh S. 2008. Repression of gene expression by unphosphorylated NF-κappaB p65 through epigenetic mechanisms. *Genes Dev*. 22: 1159-1173.

Dynan W. S. and Tjian R. 1983a. Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. *Cell* 32: 669-680.

Dynan W. S. and Tjian R. 1983b. The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell* 35: 79-87.

Fisenbeis C. F., Singh H., and Storb U. 1995. Pip, a novel IRF family member, is a lymphoid-specific, PU. 1-dependent transcriptional activator. *Genes Dev*. 9: 1377-1387.

Elbashir S. M., Lendeckel W., and Tuschl T. 2001a. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev*. 15: 188-200.

Elbashir S. M., Harborth J., Lendeckel W., Yalcin A., Weber K., and Tuschl T. 2001b. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411: 494-498.

Elbashir S. M., Harborth J., Weber K., and Tuschl T. 2002. Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* 26: 199-213.

Ernst P., Hahm K., and Smale S. T. 1993. Both LyF-l and an Eta protein interact with a critical promoter element in the murine terminal transferase gene. *Mol. Cell. Biol*. 13: 2982-2992.

Ernst P., Hahm K., Trinh L., Davis J. N., Roussel M. F., Turck C. W., and Smale S. T. 1996. A potential role for Elf-i in terminal transferase gene regulation. *Mol. Cell. Biol*. 16: 6121-6131.

Fiering S., Northrop J. P., Nolan G. P., Mattila P. S., Crabtree G. R., and Herzenberg L. A. 1990. Single cell assay of a transcription factor reveals a threshold in transcription activated by signals emanating from the T-cell antigen receptor. *Genes Dev*. 4: 1823-1834.

Fire A., Xu S., Montgomery M. K., Kostas S. A., Driver S. E., and Mello C. C. 1998. Potent and specific genetic interference by double- stranded RNA in Caenorhabditis elegans. *Nature* 391: 806-811.

Folger K., Thomas K., and Capecchi M. R. 1984. Analysis of homologous recombination in cultured mammalian cells. Cold Spring Harbor Symp. *Quant. Biol*. 49: 123-138.

Georgopoulos K., Moore D. D., and Derfier B. 1992. Ikaros, an early lymphoid-specific transcription factor and a putative mediator for T cell commitment. *Science* 258: 808-812.

Gillemans N., Tewari R., Lindeboom F, Rottier R., de Wit T., Wijgerde M., Grosveld F., and Philipsen S. 1998. Altered DNA-binding specificity mutants of EKLF and Sp1 show that EKLF is an activator of the β-globin locus control region in viva. *Genes Dev*. 12: 2863-2873.

Gilmour D. S. and Lis J. T. 1984. Detecting protein-DNA interactions in vivo: Distribution of RNA polymerase on specific bacterial genes. *Proc. Natl. Acad. Sci*. 81: 4275-4279.

Gilmour D. S. and Lis J. T. 1985. In vivo interactions of RNA polymerase II with genes of Drosophila melanogaster. *Mol. Cell. Biol*. 5: 2009-2018.

Gilmour D. S. and Lis J. T. 1986. RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells. *Mol. Cell. Biol.* 6：3984-3989.

Gu H.，Marth J. D.，Orban P. C.，Mossmann H.，and Rajewsky K. 1994. Deletion of a DNA polymerase B gene segment in T cells using cell type-specific gene targeting. *Science* 265：103-106.

Hahm K.，Ernst P.，Lo K.，Kim G. S.，Turck C.，and Smale S. T. 1994. The lymphoid transcription factor LyF-l is encoded by specific，alternatively-spliced mRNAs derived from the Ikaros gene. *Mol. Cell. Biol.* 14：7111-7123.

He L.，Hannon G. J. 2004. MicroRNAs：Small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* 5：522-531.

Hecht A.，Grunstein M. 1999. Mapping DNA interaction sites of chromosomal proteins using immunoprecipitation and polymerase chain reaction. *Methods Enzymol.* 304：399-414.

Hecht A.，Strahl-Bolsinger S.，and Grunstein M. 1996. Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. *Nature* 383：92-96.

Hollenhorst P. C.，Jones D. A.，and Graves B. J. 2004. Expression profiles frame the promoter specificity dilemma of the ETS family of transcription factors. *Nucleic Acids Res.* 32：5693-5702.

Hollenhorst P. C.，Shah A. A.，Hopkins C.，and Graves B. J. 2007. Genomewide analyses reveal properties of redundant and specific promoter occupancy within the ETS gene family. *Genes Dev.* 21：1882-1894.

Jonsen M. D.，Petersen J. M.，Xu Q. P.，and Graves B. J. 1996. Characterization of the cooperative function of inhibitory sequences in Ets-l. *Mol. Cell. Biol.* 16：2065-2073.

Kadonaga J. T.，Carner K. R.，Masiarz F. R.，and Tjian R. 1987. Isolation of cDNA encoding transcription factor Spl and functional analysis of the DNA binding domain. *Cell* 51：1079-1090.

Kim T. K. and Maniatis M. 1997. The mechanism of transcriptional synergy of an in vitro assembled interferon-β enhanceosome. *Mol. Cell* 1：119-129.

Kim D. H. and Rossi J. J. 2007. Strategies for silencing human disease using RNA interference. *Nat. Rev. Genet.* 8：173-184.

Krylov D.，Kasai K.，Echlin D. R.，Taparowsky E. J.，Arnheiter H.，and Vinson C. 1997. A general method to design dominant negatives to B-HLHZip proteins that abolish DNA binding. *Proc. Natl. Acad. Sci.* 94：12274-12279.

Langlands，K.，Yin X.，Anand G.，and Prochownik E. V. 1997. Differential interactions of Id proteins with basic helix-loop- helix transcription factors. *J. Biol. Chem.* 272：19785-19793.

Lassus P.，Rodriguez J.，and Lazebnik Y. 2002. Confirming specificity of RNAi in mammalian cells. *Sci. STKE* 147：PL13.

Lauring J. and Schlissel M. S. 1999. Distinct factors regulate the murine RAG-2 promoter in B and T cell lines. *Mol. Cell. Biol.* 19：2601-2612.

Lee G. M.，Donaldson L. W.，Pufall M. A.，Kang H. S.，Pot I.，Graves B. J.，and Mcintosh L. P. 2005. The structural and dynamic basis of Ets1 DNA binding autoinhibition. *J. Biol. Chem.* 280：7088-7099.

Lefstin J. A. and Yamamoto K. R. 1998. Allosteric effects of DNA on transcriptional regulators. *Nature* 392：885-888.

Lehman A. M.，Ellwood K. B.，Middleton B. E.，and Carey M. 1998. Compensatory energetic relationships between upstream activators and the RNA polymerase II general transcription machinery. *J. Biol. Chem.* 273：932-939.

Li Q. and Johnston S. A. 2001. Are all DNA binding and transcription regulation by an activator physiologically relevant? *Mol. Cell. Biol.* 21：2467-2474.

Lipston K. E. and Baserga R. 1989. Transcriptional activity of the human thymidine kinase gene determined by a method using the polymerase chain reaction and an intron-specific probe. *Proc. Natl. Acad. Sci.* 86：9774-9777.

Lo K., Landau N. R., and Smale S. T. 1991. LyF-1, a transcriptional regulator that interacts with a novel class of promoters for lymphocyte-specific genes. *Mol. Cell. Biol.* 11: 5229-5243.

Lodish H., Baltimore D., Berk A., Zipursky S. L., and Darnell J. 1996. *Molecular cell biology*, 3rd ed. W. H. Freeman, New York.

Martone R., Euskirchen G., Bertone P., Hartman S., Royce T. E., Luscombe N. M., Rinn J. L., Nelson F. K., Miller P., Gerstein M., et al. 2003. Distribution of NF-κB-binding sites across human chromosome 22. *Proc. Natl. Acad. Sci.* 100: 12247-12252.

Mauxion F., Jamieson C., Yoshida M., Arai K., and Sen R. 1991. Comparison of constitutive and inducible transcriptional enhancement mediated by κB-related sequences: Modulation of activity in B cells by human T-cell leukemia virus type I tax gene. *Proc. Natl. Acad. Sci.* 88: 2141-2145.

Miyagishi M. and Taira K. 2002. U6 promoter-driven siRNAs with four uridine 3′ overhangs efficiently suppress targeted gene expression in mammalian cells. *Nat. Biotechnol.* 20: 497-500.

Okamura R. M., Sigvardsson M., Galceran J., Verbeek S., Clevers H., and Grosschedl R. 1998. Redundant regulation of T cell differentiation and TCRce gene expression by the transcription factors LEF-l and TCF-1. *Immunity* 8: 11-20.

Olive M., Williams S. C., Dezan C., Johnson P. F., and Vinson C. 1996. Design of a C/EBP-specific, dominant-negative bZIP protein with both inhibitory and gain-of-function properties. *J. Biol. Chem.* 271: 2040-2047.

Orlando V. and Paro R. 1993. Mapping Polycomb-repressed domains in the bithorax complex using in vivo formaldehyde cross-linked chromatin. *Cell* 75: 1187-1198.

Orlando V., Strult H., and Pam R. 1997. Analysis of chromatin structure by in vivo formaldehyde cross-linking. *Methods* 11: 205-214.

Paddison P. J. 2008. RNA interference in mammalian cell systems. *Curr. Top. Microbiol. Immunol.* 320: 1-19.

Paddison P. I., Caudy A. A., Bernstein E., Hannon G. J., and Conklin D. S. 2002. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.* 16: 948-958.

Panne D., Maniatis T., and Harrison S. C. 2007. An atomic model of the interferon-β enhanceosome. *Cell* 129: 1111-1123.

Parekh B. S. and Maniatis T. 1999. Virus infection leads to localized hyperacetylation of histones H3 and H4 at the IFN-β promoter. *Mol. Cell* 3: 125-129.

Paul C. P., Good P. D., Winer I., and Engelke D. R. 2002. Effective expression of small interfering RNA in human cells. *Nat. Biotechnol.* 20: 505-508.

Pe Y., and Tuschl T. 2006. On the art of identifying effective and specific siRNAs. *Nat. Methods* 3: 670-676.

Phuc Le P., Friedman J. R., Schug J., Brestelli J. E., Parker J. B., Bochkis I. M., and Kaestner K. H. 2005. Glucocorticoid receptor-dependent gene regulatory networks. *PLoS Genet.* 1: e16.

Plevy S. F., Gemberling J., Hsu S., Dorner A. J., and Smale S. T. 1997. lvlultiple control elements mediate activation of the murine and human IL-12 p40 promoters: Evidence of functional synergy between C/EBP and Rd proteins. *Mol. Cell. Biol.* 17: 4572-4588.

Pongubala J. M., Nagulapalli S., Klernsz M. J., McKercher S. R., Maki R. A., and Atchison M. L. 1992. PU. 1 recruits a second nuclear factor to a site important for immunoglobulin K 3′ enhancer activity. *Mol. Cell. Biol.* 12: 368-378.

Pufall M. A., Lee G. M., Nelson M. L., Kang H. S., Velyvis A., Kay L. E., Mcintosh L. P., and Graves B. J. 2005. Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region. *Science* 309: 142-145.

Reddy S. T., Gilbert R. S., Xie W., Luner S., and Herschman H. R. 1994. TGF-β inhibits both endotoxin-induced prostaglandin synthesis snd expression of the TIS10/prostaglandin synthase 2 gene in murine macrophages. *J. Leukoc. Biol.* 55: 192-200.

Rishi V. and Vinson C. 2003. Dominant-negative mutants of helix- loop-helix proteins: Transcriptional inhibition. *Methods Enzymol*. 370: 454-466.

Rundlett S. E., Carmen A. A., Suka N., Turner B. M., and Grunstein M. 1998. Transcriptional repression by UME6 involves deacetylation of lysine 5 of histone H4 by RPD3. *Nature* 392: 831-835.

Sen G. L. and Blau H. M. 2006. A brief history of RNAi: The silence of the genes. *FASEB J*. 20: 1293-1299.

Shah P. C., Bertolino F., and Singh H. 1997. Using altered specificity Oct-l and Oct-2 mutants to analyze the regulation of immunoglobulin gene transcription. *EMBO J*. 16: 7105-7117.

Smithies O., Gregg R. G., Boggs S. S., Koralewski M. A., and Kucherlapati R. S. 1985. Insertion of DNA sequences into the human chromosomal β-globin locus by homologous recombination. *Nature* 317: 230-234.

Snóve O., Jr. and Rossi J. J. 2006. Expressing short hairpin RNAs in vivo. *Nat. Methods* 3: 689-695.

Solomon M. J. and Varshavsky A. 1985. Formaldehyde-mediated DNA-protein crosslinking: A probe for in vivo chromatin structures. *Proc. Natl. Acad. Sci.* 82: 6470-6474.

Solomon M. J., Larsen P. L., and Varshavsky A. 1988 . Mapping protein- DNA interactions in vivo with formaldehyde: Evidence that histone H4 is retained on a highly transcribed gene. *Cell* 53: 937-947.

Stein C. A. 1998. How to design an antisense oligodeoxynucleotide experiment: A consensus approach. *Antisense Nucleic Acid Drug Dev*. 8: 129-132.

Svoboda P. 2007. Off-targeting and other non-specific effects of RNAi experiments in mammalian cells. *Curr. Opin. Mol. Ther*. 9: 248-257.

Thanos D. and Maniatis T. 1995a. Virus induction of human IFN β gene expression requires the assembly of an enhanceosome. *Cell* 83: 1091-1100.

Thanos D. and Maniatis I. 1995b. NF-κB: A lesson in family values. *Cell* 80: 529-532.

Thomas K. R. and Capecchi M. R. 1987. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* 51: 503-512.

van Genderen C., Okamura R. M., Farinas I., Quo R. G., Parslow T. G., Bruhn L., and Grosschedl R. 1994. Development of several organs that require inductive epithelial-mesenchymal interactions is impaired in LEF-1-deficient mice. *Genes Dev*. 8: 2691-2703.

Verbeek S., Izon D., Hofhuis F., Robanus-Maandag E., te Ride H., van de Wetering M., Oosterwegel M., Wilson A., MacDonald H. R., and Clevers H. 1995. An HMG-box-containing T-cell factor required for thymocyte differentiation. *Nature* 374: 70-74.

Walter J. Dever C. A., and Biggin M. D. 1994. Two homeo domain proteins bind with similar specificity to a wide range of DNA sites in Drosophila embryos. *Genes Dev*. 8: 1678-1692.

Wathelet M. G., Lin C. H., Parekh B. S., Ronco L. V., Howley P. M., and Maniatis T. 1998. Virus infection induces the assembly of coordinately activated transcription factors on the IFN-β enhancer in vivo. *Mol. Cell* 1: 507-518.

Wiznerowicz M., Szulc J., and Trono D. 2006. Tuning silence: Conditional systems for RNA interference. *Nat. Methods* 3: 682-688.

9 内源性控制区的体内分析

要点

- 染色质免疫沉淀（ChIP）技术是检查特定因子和被修饰组蛋白与目的内源性 DNA 区域结合，以及研究全基因组范围内蛋白质-DNA 相互作用最普遍的方法。但是，ChIP 也具有不容忽视的局限性。
- 检查蛋白质-DNA 相互作用的替代策略包括 DamID 及使用 DNase I、硫酸二甲酯和高锰酸钾的基因组足迹法，这些策略有时可以克服 ChIP 的局限性。
- 可以使用各种各样的实验技术监测核小体在单个基因座或全基因组水平上的密度和定位，这些实验技术大部分依赖于微球菌核酸酶对无核小体区及核小体连接区的特异性。
- 通常通过使用核酸酶可及性分析监测核小体重塑和染色质结构中的动态变化。
- 染色体内及染色体间相互作用对于正确转录调控的重要性，已开始从使用染色质构象捕获（3C）分析的研究中显现出来。
- 转录控制区的胞嘧啶甲基化状态可在基因表达的激活和沉默中发挥重要作用，并且可以在单个基因或全基因组上进行监测。

引言和概述

前面的章节描述了鉴定目的基因重要控制区和表征相关蛋白质-DNA 相互作用的策略。然而，由于通常用于研究转录调控的很多方法的人工性质和局限性，很难确定哪个转录因子在其自然基因组位置与控制区相互作用。此外，现在已广泛地意识到染色质在转录调控中起着重要作用，这使得有必要表征处于内源性基因座的染色质结构，以实现对调控机制的完全理解。

我们已描述了许多实验策略，它们能够提供关于基因在其自然基因组位置转录的有价值信息。例如，第 3 章中转录起始位点定位技术揭示内源性基因上转录起始的位置。在第 4 章和第 5 章中，我们讨论了细菌人工染色体（BAC）转基因和靶向敲入的使用，作为在天然或近天然染色体环境中研究控制区的功能性方法。第 5 章也包括了对 DNase I 超敏性、染色质免疫沉淀（ChIP）及用于鉴定远程控制区的系统发生策略的讨论。最后，第 8 章描述了 ChIP 的使用、转录因子功能缺失策略，以及用于检测目的转录因子对候选靶基因调控的相关性的其他策略。

序列特异性蛋白质-DNA相互作用

- 染色质免疫沉淀（ChIP）
- DamID
- DNase I基因组足迹
- DMS基因组足迹

DNA熔解

- 高锰酸钾基因组足迹

核小体定位

- MNase-Southern印迹
- MNase-PCR扫描
- MNase-LM-PCR
- MNase引物延伸
- MNase-微阵列
- MNase-测序

核小体重塑

- DNase I敏感性(第5章)
- DNase I高敏感性(第5章)
- MNase方法
- 限制酶可及性

染色体内和染色体间的相互作用

- 染色质构象捕获(3C和相关方法)

DNA甲基化

- 甲基化敏感性限制酶分析
- 亚硫酸氢盐测序
- 限制性标记基因组扫描(RLGS)
- 甲基化DNA免疫沉淀(mDIP)

图 9.1 分析内源性控制区的方法。

在这一章，我们更深入地描述许多可用于表征内源性基因座的实验策略（图9.1）。本章的前半部分聚焦于可用于监测在一个基因组位置上蛋白质-DNA 相互作用的技术，这些相互作用包括特定转录因子与辅因子的相互作用及修饰组蛋白与其他染色质蛋白的相互作用。在这一讨论中包括 ChIP（同样见第 5 章和第 8 章）、DamID 和使用 DNase I、硫酸二甲酯（DMS）及高锰酸盐的基因组足迹。本章的后半部分探索用于进一步表征染色质结构和一个基因组座位其他特征的策略。我们概述了用于监测核小体定位、核小体重塑、染色体内和染色体间的相互作用及 DNA 甲基化的策略。在这一章的结尾，介绍了少数几个相关的技术，但在大多数情况下，读者可参考别处出版的方案。

对于对目的基因调控机制的深入理解感兴趣的研究者来说，本章讨论的策略非常有价值。然而，在继续之前，重要的是要强调这些技术不是功能性技术，所获得的结果将提供关于染色质结构和与目的控制区结合的蛋白质的有价值的信息。然而，为了确定观察到的蛋白质和染色质特征是否真正对转录调控重要，同时为了确定它们如何促成正确调控，最终将需要功能性研究，这些功能性研究通常类似于第 8 章中讨论的那些技术。

实验策略

染色质免疫沉淀

在前面的章节中，作为一种研究特定蛋白质或修饰型蛋白质与基因组 DNA 区域相互作用的宝贵方法，已经介绍了 ChIP。第 8 章（见图 8.1）中提供了该实验方法和该方法的流程图，并且在本章末尾（见方案 9.3）也包括了一个 ChIP 的基本实验方案。虽然第 8 章中的讨论主要聚焦于使用 ChIP 确定一个转录因子是否与候选靶基因相互作用，但是 ChIP 也以相同的频率用于监测翻译后修饰组蛋白在特定基因组位置的出现。

ChIP 还经常被用于在全基因组规模上鉴定转录因子和组蛋白的位置（Ren et al. 2000；Weinmann et al. 2002；Hanlon and Lieb 2004；Elnitski et al. 2006；Kim and Ren 2006）。在大多数早期研究中，首先纯化免疫沉淀的 DNA 片段，再通过全基因组扩增、连接介导聚合酶链反应（LM-PCR）或其他方法（见 O' Geen et al. 2006）扩增纯化的 DNA 片段。之后，将扩增的 DNA 片段与叠连 DNA 阵列退火，其中所使用的叠连 DNA 阵列覆盖整个基因组、特定的染色体，或大量假定启动子片段。这些 ChIP-chip（或 ChIP-on-chip）方法的详细实验策略和程序可以在合成叠连阵列的公司的网站上找到，包括 Agilent、NimbleGen 和 Affymetrix。最近，ChIP-测序（ChIP-sequencing，ChIP-Seq）方法已经取代 ChIP-chip 成为选择的方法，ChIP-Seq 技术主要程序包括解交联交联后免疫沉淀的 DNA、DNA 片段化，以及大量并行的 DNA 测序分析（Mardis 2007，2008；Barski et al. 2007；Robertson et al. 2007；Jothi et al. 2008；Wang et al. 2008）。ChIP-Seq 的主要优点是，该技术能提供更加精确和定量的信息。不像 ChIP-chip，ChIP-Seq 不受基因组 DNA 片段与固定在叠连 ChIP-chip 微阵列上的寡核苷酸的高度易变且通常弱的退火效率的影响。同时，据认为 ChIP-Seq 可获得比

ChIP-chip 分辨率更高的数据，因为它提供沉淀片段的实际 DNA 序列。

ChIP、ChIP-chip 和 ChIP-Seq 方法能够提供与个别目的控制区联系的蛋白质的宝贵信息，或者与转录因子、染色质蛋白或具有特定翻译后修饰的组蛋白相联系的基因组位点集合的宝贵信息。这些结果通常对发展特定转录因子或组蛋白修饰功能相关的假说非常有用。相反，这些方法可用于验证使用其他策略产生的假说，如一个特定转录因子与在特定发育阶段或对特定刺激应答中表达的基因的控制区直接而稳定联系的假说。此外，正如第 2 章中所提到的，ChIP-chip 和 ChIP-Seq 的结果越来越频繁地用于建立转录网络（Blais and Dynlacht 2005）。

尽管 ChIP 方法有巨大的价值，但也务必要意识到它们的局限性。第 8 章中已讨论了其中部分局限性，尤其是，第 8 章强调了：①与阴性对照相比，ChIP 分析通常产生低信号，这会导致非决定性的结果；②由于该分析的分辨率有限，难以确定一个因子精确的结合位点；③ChIP 不是功能性分析方法，并且单独的 ChIP 分析不能表明发现定位于目的基因组区域的蛋白质或修饰型组蛋白的功能意义。这最后一点局限性是尤其成问题的，因为许多研究现在已提供了有力的证据，证明蛋白质与基因组位点结合，但蛋白质在这些位点上却没有明显的功能（Walter et al. 1994；Li and Johnston 2001；Martone et al. 2003；Phuc Le et al. 2005；Beima et al. 2006；Hollenhorst et al. 2007；Dong et al. 2008；Li et al. 2008）。一直难以设计实验来严格地测验这种可能性，即相对于转录因子-DNA 相互作用，特定组蛋白修饰在它们被发现的特定基因组位置上不具有功能重要性。不过，考虑到以转录因子观察到的无功能相互作用，重要的是要牢记这种可能性，即被修饰组蛋白存在在基因组位置上，但在这些位置上该修饰没有功能意义。

ChIP 分析的另外一个局限性是，即使在免疫沉淀和 Western 印迹实验中似乎高度特异的抗体，也经常有可能与其他细胞核蛋白交叉反应。一项可以揭示交叉反应活性但却很少做的实验是，在缺少抗体靶向的特定因子的细胞中进行平行 ChIP 分析。Beima 等进行的转录因子 T-bet（T 细胞中表达的 T-box）靶标的分析提供了一个这种对照实验的例子，该 ChIP-chip 研究鉴定了新的 T-bet 靶基因，并使用许多实验策略对这些新的靶基因进行了仔细评估。有一项实验是监测在 T-bet 缺陷细胞中当使用 T-bet 抗体时在候选靶基因获得的 ChIP 信号。令人惊讶的是，在最初的 ChIP-chip 分析中，以及在以单个 PCR 引物对进行的后续 ChIP 实验中，表现出强烈信号的候选位点，其中约 10%在 T-bet 缺陷细胞中产生了类似幅度的阳性结果。这一结果显示，T-bet 抗体虽然在标准尺度上高度特异，但仍与其他细胞核蛋白交叉反应。因此，即使是最强的 ChIP 信号也必须谨慎解释。

对于 ChIP-chip 和 ChIP-Seq 研究来说，与低信号和低信噪比相关的困难被进一步扩大。只要有可能，就应使用传统 ChIP 分析测试在全基因组研究中获得的结果的有效性，这非常重要。对于 ChIP-chip 和 ChIP-Seq 结果来说，另一个重要的问题是，大量的数据，加上基因调控策略的复杂性，使得其难以得出有意义的结论。在文献中已出现这样一种倾向，即基于已经观察到的趋势得出肯定的结论，即使偏离和符合这一趋势的基因一样多。关于 ChIP 实验的设计和解释的其他技术信息，可以在方案 9.3 中找到，

也可以在专门从事 ChIP 实验抗体生产的公司的网站上找到，如 Abcam，还可以在发表的文献中可以找到（Bernstein et al. 2004；Buck and Lieb 2004；Chaya and Zaret 2004；Ciccone et al. 2004；Das et al. 2004；Ren and Dynlacht 2004；Elnitski et al. 2006；Kim and Ren 2006；Southall and Brand 2007）。

DamID

在上面举例的 ChIP 局限性中，不与其他的蛋白质交叉反应的高质量抗体的需求，以及由低效率交联或转录因子瞬时结合引起的弱信号的频繁出现，是最显著的两个局限性。DamID（DNA 腺嘌呤甲基转移酶识别）是由 Van Steensel 和 Henikoff（2000）发展的一种分析方法，该分析已在数项研究中表明有潜力避免这些局限性（Southall and Brand 2007；Vogel et al. 2007；Zhang et al. 2007；Venkatasubrahmanyam et al. 2007；Reddy et al. 2008）。事实上，考虑到这一技术在数项已发表的研究中取得的成功，可以说对于那些在 ChIP 分析中产生不佳结果的蛋白质来说，这一技术严重未充分利用，因为在 PubMed 中使用 DamID 作为关键词搜索时，得到的结果不足 20 篇。在最近的一项研究中，没有先前 DamID 使用经验的实验室提供了强有力的证据，表明在鉴定酵母 Sir 蛋白相互作用位点的能力上，DamID 优于 ChIP（Venkatasubrahmanyam et al. 2007）。

DamID 分析始于一个融合蛋白，即目的转录因子或染色质蛋白与大肠杆菌 DNA 腺嘌呤甲基转移酶（Dam）之间的融合蛋白。Dam 甲基化 GATC 序列上腺嘌呤的 N6 位置。当在动物或真核细胞系中表达了融合蛋白时，Dam 将优先甲基化其融合伙伴的基因组结合位点附近的腺嘌呤。然后，可通过多种策略鉴定腺嘌呤甲基化的基因组位点。在 Van Steensel 和 Henikoff（2000）的原创性研究中，用限制性内切核酸酶 *Dpn*II 切割基因组 DNA，该酶只切割未甲基化的 GATC 模体，而不切割带有甲基腺嘌呤的 GATC 模体。之后，以位于目的区域中个别 GATC 侧翼的引物进行定量 PCR 检测。

最新的研究已经使用了全基因组的方法（图 9.2；Vogel et al. 2007）。首先将甲基化的基因组 DNA 用限制性内切核酸酶 *Dpn*I 切割，该酶特异性地在甲基化 GATC 模体的 A 和 T 之间切割。因为靶向于某一特定位点的 Dam 融合蛋白催化的甲基化往往发生在一个相对宽的窗口（见 Van Steensel and Henikoff 2000），并且由于在多个甲基化 GATC 模体上的切割，*Dpn*I 的切割通常导致在融合蛋白结合位点附近产生 DNA 小片段。接下来，将带有一个 TC 序列开头平末端的双链接头与切割的 DNA 混合物连接。由于基因组 DNA 片段以 GA 结尾，并且连接到由 TC 序列开始的接头上，因此连接反应再造了 GATC。连接之后，用 *Dpn*II 孵育 DNA，*Dpn*II 切割未甲基化 GATC 模体，因此也就消除了任何缺少期望的 $G^{me6}ATC$ 的连接产物。然后，以退火至接头序列的引物进行 PCR，扩增连接了接头的 DNA 片段，再将扩增片段与基因组叠连阵列退火。

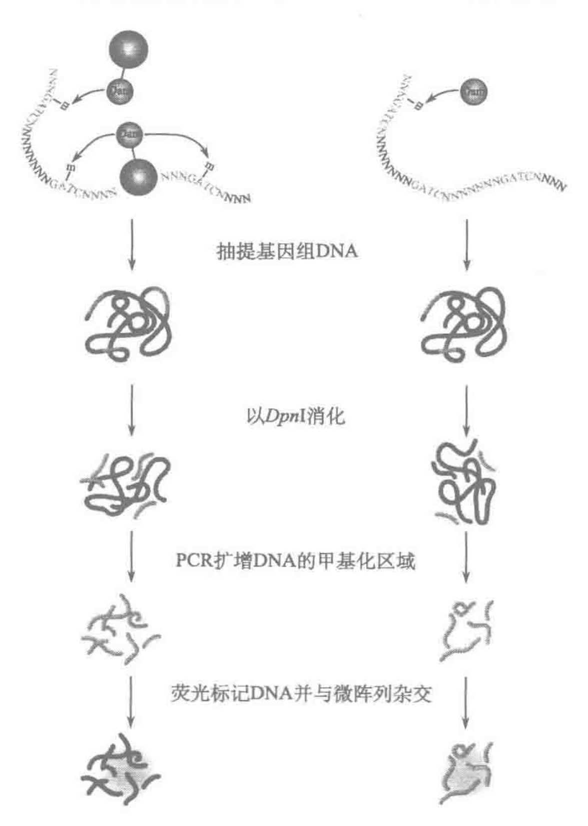


图 9.2　作为一种鉴定目的因子基因组相互作用位点的 DamID 方法（经 Oxford University Press 允许，重绘自 Southall and Brand 2007）。

与 ChIP 相比，DamID 的优点是它的灵敏性及其无需抗体和交联的事实，因此消除了由特异性和亲和性不足的抗体引起的问题，也避免了由于交联引起的实验假象。然而，DamID 也有其自身的局限性。

1. Dam 是一种高活性酶，可以有效地、无特异性地靶向甲基化基因组 DNA 中的腺嘌呤。因此，重要的是要在异常低的水平上表达融合蛋白，使无目标的甲基化最小化；同时以缺少融合伙伴的 Dam 进行平行对照实验，监测背景信号。

2. Dam 基团可能会改变融合伙伴的活性，从而造成融合伙伴与某些基因组位置的结合，而这些基因组位置却不同于内源性蛋白结合的基因组位置。

3. 与 ChIP 不同，DamID 不能用于研究具有特定翻译后修饰的蛋白质的基因组结合，如修饰的组蛋白。

4. 因为在 Dam 定位到一个特定的基因位置之后，它造成的甲基化可跨越数百个碱基对，所以该分析的分辨率甚至低于 ChIP。

5. 虽然 DamID 能提供与目的因子结合的广泛 DNA 区域的有价值信息，但是它却

几乎不提供该广泛区域中特定结合位点的信息。

尽管存在这些局限性，对于许多重要的但不适合由 ChIP 分析的转录因子和染色质蛋白，DamID 应视为一个很有吸引力的选择。

DNase I 和 DMS 基因组足迹

DNase I 和 DMS 基因组足迹是用于检测内源控制区上蛋白质-DNA 相互作用的方法（Mueller et al. 1988；Mueller and Wold 1989；Garrity and Wold 1992；Garrity et al. 1995）。ChIP 和 DamID 提供关于可结合控制区的蛋白质的身份信息，但由于它们的低分辨率，这些分析不能揭示精确的相互作用位点。相反，基因组足迹不能揭示蛋白质的身份，但它可以提供关于蛋白质结合的控制区的特定 DNA 模体的高分辨率信息。基因组足迹提供的信息是无偏的，然而 ChIP 和 DamID 分析只能提供为分析而专门选定的蛋白质的相关信息。基因组足迹法能揭示使用其他方法不能检测的重要蛋白质-DNA 相互作用，并且可帮助确证相互作用的重要性，这些程序和本章中所描述的其他程序依赖于被称为 LM-PCR 的技术（见专题 9.1）。LM-PCR 可与切割或修饰完整细胞或细胞核中的基因组 DNA 的各种试剂联用，其中每种试剂提供关于基因调控的独特认识。

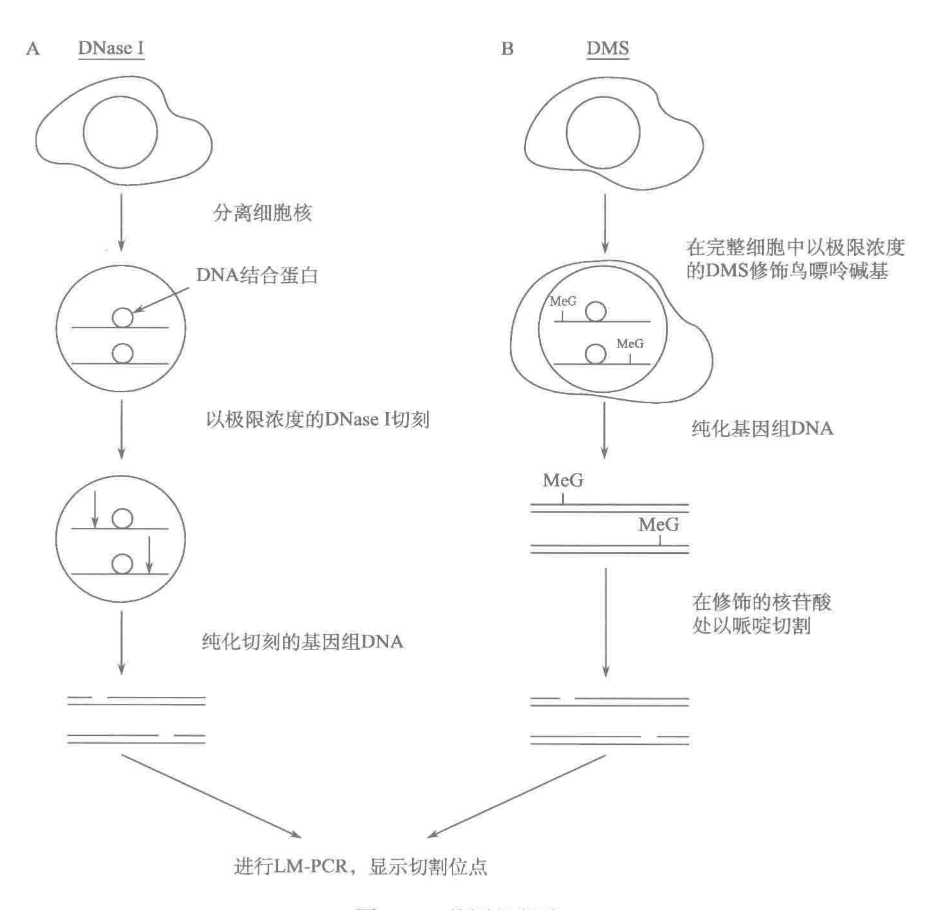


图 9.3 基因组足迹。

为了进行基因组足迹实验，可在分离的细胞核中用 DNase I 切割基因组 DNA（图 9.3A 和方案 9.2）。然后，纯化切割的基因组 DNA 并通过 LM-PCR 分析，或者可通过 DMS（图 9.3B 和方案 9.2；也见第 11 章）对完整细胞内的基因组 DNA 在鸟苷和腺苷残基上进行修饰。然后，纯化修饰的 DNA，再用哌啶处理，催化修饰碱基邻近的水解，最后通过 LM-PCR 分析 DNA。对于 DNase I 和 DMS 这两种方法，都需要对照反应；在对照反应中，以相同的程序分析纯化的基因组 DNA 或含该基因座的基因组克隆。这提供了非结合型 DNA 的切割模式，以便与在细胞中经 DMS 修饰或在细胞核中经 DNase I 切割后观察到的模式进行比较。蛋白质-DNA 相互作用显示为保护的区域和（或）高敏感性（关于足迹理论的基本讨论，见第 7 章和第 11 章）。

专题 9.1 连接介导 PCR

由于其极高的敏感性和特异性，LM-PCR 成为检测复杂样品中 DNA 链断裂的一项强大技术。Barbara Wold 及其同事（Mueller et al. 1988；Mueller and Wold 1989）首次描述了 LM-PCR（图 9.4），同一实验室随后对该技术进行了改进（Garrity and Wold

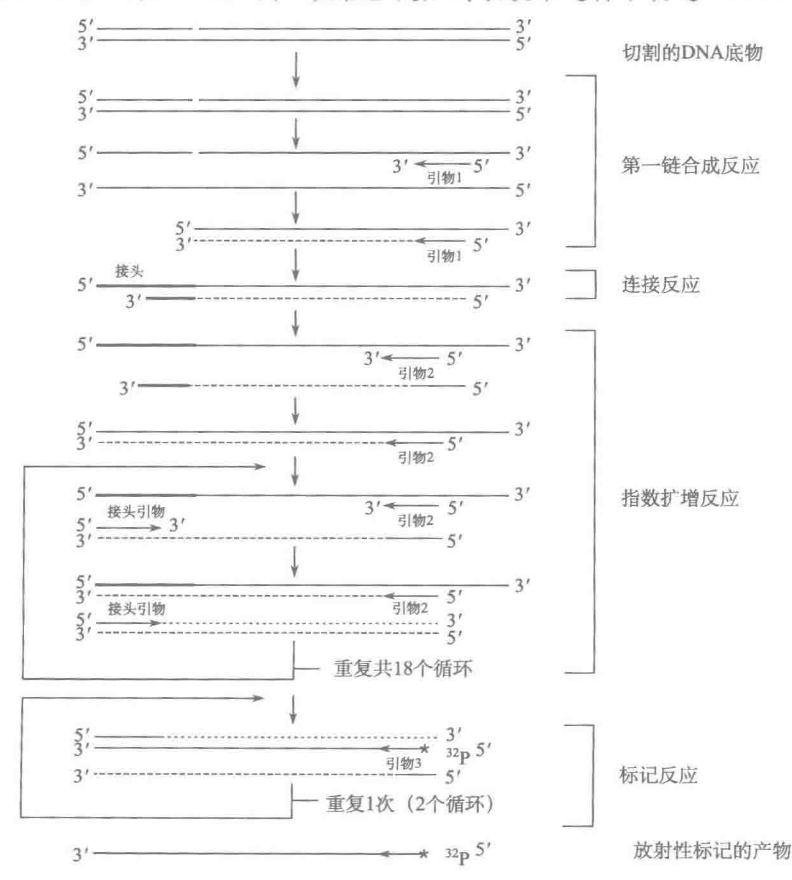


图 9.4 LM-PCR（经许可，改编自 Garrity and Wold 1992。）

1992；Garrity et al. 1995；也见方案 9.2)。这项技术一般在纯化的基因组或质粒 DNA 上进行，其中，基因组或质粒 DNA 之前已经用极限浓度的核酸酶（如 DNase I、微球菌核酸酶或限制性内切核酸酶）进行了切割（图 9.4)。或者，可以在已经以 DMS 或高锰酸钾修饰的 DNA 上进行 LM-PCR，随后通过哌啶在被修饰的核苷酸上进行 DNA 链的切割。

在大多数 LM-PCR 程序中，第一个步骤是将合成的寡核苷酸引物（引物 1）与经核酸酶或哌啶切割并变性的 DNA 的特定序列退火。引物 1 通常退火在距目的区域至少 100 bp 处（见方案 9.2)。用热稳定性 DNA 聚合酶将引物分子延伸到末端（即切割位点)，产生平末端。然后，将寡核苷酸接头与平末端的 DNA 片段连接。这一步骤中所用的接头含有一个平末端和一个黏性末端，以便总是以相同方向连接（注意：用 MNase 切割后，LM-PCR 程序的第一个步骤会有所不同；见专题 9.3 和方案 9.2)。由于接头序列是已知的，故可以设计引物（称为接头引物)，使其与 PCR 中使用的序列退火。

PCR 的第一步以接头引物和引物 2 进行。引物 2 通常杂交到与引物 1 序列重叠的序列上，但比引物 1 序列稍微靠内，从而确保只有那些来自目的基因的 DNA 片段被扩增。对于基因组足迹程序来说，这一扩增步骤产生一套嵌套的双链 DNA 片段。第一个扩增步骤之后，使用同样的接头引物和比引物 2 稍靠内的另一个引物 3 进行第二次 PCR。使用前，引物 3 使用 T4 多核苷酸激酶和 [γ-^{32}P] ATP 进行 5′端标记。这一最终 PCR 导致放射性标记 DNA 片段的扩增，其中 DNA 片段的大小对应于从引物 3 的 5′端到核酸酶或化学药品切割位点的距离，并加上接头引物的长度。由于起始点是一套嵌套的 DNA 片段，最后的放射性标记产物应该跨越一个宽的尺寸范围。这些片段可以在变性聚丙烯酰胺凝胶上分析，随后进行放射自显影或磷屏成像分析。

与体外 DNase I 足迹相比，使用 DNase I 的基因组足迹提供由于 DNA 结合蛋白的占据而不能被核酸酶消化的 DNA 序列信息（见第 7 章和第 11 章)。像类似的体外技术（见第 11 章）一样，DMS 保护技术提供由于结合蛋白质的紧密结合而不能被 DMS 修饰的核苷酸的信息。因为 DMS 是一种小分子化学药品，受结合蛋白质保护而未被修饰的 DNA 区域远远短于在 DNase I 消化中被保护的 DNA 区域，并且可能与蛋白质结合所需的核苷酸更为一致（见第 11 章)。

基因组足迹的应用和局限

以 DNase I 或 DMS 进行的基因组足迹分析可作为剖析控制区的出发点，因为它能洞悉功能性重要 DNA 元件的位置。然而，它通常不能取代系统进化分析或综合性突变分析（见第 6 章)，因为只有一小部分重要控制元件才有可能产生基因组足迹。而且，在关于相关结合蛋白的身份，以及相对于调控同一控制区的其他蛋白质，该结合蛋白对基因调控的重要性方面，基因组足迹只提供有限的信息。正如第 8 章所讨论的，基因组足迹可以为候选调控蛋白的功能相关性提供支持。如果在内源控制元件上观察到的基因组足迹模式与在体外使用纯化的重组蛋白质观察到的足迹模式一致（当两个反应一起进行时)，则说明该蛋白质可能确实负责内源元件的功能。然而，对此实验的一项告诫是，

如果体内 DNA 组装成核小体，则在体内观察到的基本切割模式与体外观察到的切割模式会有很大的差别；核小体 DNA 会展示不同模式的切割产物，从而使体内和体外足迹之间的比较难以进行。基因组足迹的第三个基本应用是确定内源性控制区中 DNA 元件的占据是否是可诱导的，或者是否受细胞类型特异性的调控。

一个知之甚少的问题是为什么有些重要蛋白质-DNA 相互作用不能产生可检测的基因组足迹，可以考虑以下几个可能的原因。对于 DNase I 基因组足迹来说，一个可能是重要的蛋白质可能与一段 DNA 序列相互作用，而该段 DNA 序列在缺乏蛋白质结合的情况下，对 DNase I 切割具有固有的抗性（见第 7 章和第 11 章）。如果 DNase I 不能有效地切割无蛋白质的对照模板的一个区域，将难以检测蛋白质存在时该区域中的保护。另一种可能是在细胞核制备期间，蛋白质可能从 DNA 解离。例如，在 DMS 足迹中，DMS 可能直接修饰 DNA 结合蛋白，从而引起 DNA 结合蛋白从 DNA 解离。还有一种可能是蛋白质可能未与大沟鸟嘌呤或小沟腺嘌呤紧密接触，而这些嘌呤是 DMS 的靶点。最后一种可能是 DNase I 或 DMS 保护可能被大量的 PCR 循环所掩盖，PCR 可能优先扩增保护区域中的背景切割产物或弱切割产物。

除了上述纯粹的技术原因之外，还有其他一些具有生物学意义的原因也会导致通过基因组足迹可能无法检出蛋白质-DNA 相互作用。例如，在用于实验的细胞群中，只有一部分细胞中发生该相互作用。这对于可诱导基因来说是一个常见的问题，可诱导基因通常仅在一部分细胞中被诱导，即使当实验中使用单克隆细胞系时也是如此。*Il2* 和 *Il12b* 基因就是两个可诱导基因的例子，它们在单克隆细胞群中有时仅在一小部分细胞中被诱导（Garrity et al. 1994；Weinmann et al. 2001）。为了澄清这种可能性，可以先通过流式细胞仪或免疫荧光监测表达效率。如果这些实验确证基因仅在一小部分细胞中表达，可以考虑使用能将表达细胞与未表达细胞分离的方法［如荧光激活的细胞分选（FACS）或磁珠］。切记，除非细胞群体中至少 75%～80%的位点被蛋白质占据，否则保护区域是难以检测的（见第 7 章和第 11 章）。如果未观察到高效的表达，并且如果难以获得分离表达细胞的方法，就可能难以获得令人信服的基因组足迹结果。

通过基因组足迹可能不能检测蛋白质-DNA 相互作用的另外一个原因是，在每个细胞中蛋白质可能仅占据了两个等位基因上的一个位点，从而导致基因的单等位基因表达。近期的研究已表明，在每个细胞内大量的哺乳动物基因可能只在一个等位基因上表达（Gimeibrant et al. 2007）。然而，单等位基因表达并不一定消除基因组足迹实验成功的可能性。在某些和或许大部分单等位基因表达的例子中，大部分的转录激活因子和相关的染色质蛋白可能与两个等位基因都结合，转录之所以只发生在其中一个等位基因上，是因为全套因子驱使前起始复合物组装，并在适度比例的等位基因上起始转录（见 Smale and Fisher 2002；Guo et al. 2005）。

已发现某些转录因子，包括糖皮质激素受体和核因子 κB（NF-κB），与内源识别位点瞬时并动态地结合（McNally et al. 2000；Bosisio et al. 2006），这暗示了在基因组足迹实验中某些蛋白质可能难以检测的另一个可能的原因。与该可能性相一致的是，*Il12b* 启动子的基因组足迹揭示了与 CCAAT/增强子结合蛋白（C/EBP）位点的强相互作用，但没有检测到与 NF-κB 功能性重要结合位点的相互作用（Weinmann et al. 1999），已经表明，

NF-κB是以瞬时方式与靶位点结合的蛋白质之一（Bosisio et al. 2006)。

尽管不能检测某些蛋白质-DNA 的相互作用，但是 DNase I 和 DMS 基因组足迹能提供发生在内源控制区上的相互作用的有价值信息。当与功能性分析及特定 DNA 结合蛋白的研究相结合时，该信息可以导致对基因调控机制的更全面的理解。在本章的最后提供了 DNase I 和 DMS 基因组足迹方案（方案 9.2)。

高锰酸钾基因组足迹

高锰酸钾是这里讨论的最后一种能与 LM-PCR 联用以分析内源基因座的试剂。在第 12 章中，当讨论用于剖析转录激活机制的体外分析方法时，详细讨论了高锰酸钾足迹的理论基础。简而言之，高锰酸钾能够进入完整细胞并修饰胸腺嘧啶，从而标记这些胸腺嘧啶，以便随后的哌啶切割。然而，高锰酸钾是独特的，因为它修饰单链 DNA 中的胸腺嘧啶比修饰双链 DNA 中的胸腺嘧啶更有效。因此，解链的基因组基因座区域对这种试剂的切割是敏感的。由于这种独特的性质，高锰酸钾足迹最初用于鉴定解链的 DNA。然而，也可以观察到受保护区域，从而提供关于蛋白质-DNA 相互作用的信息，这些信息与 DNase I 和 DMS 基因组足迹法所提供的信息相似。用于后一个目的的高锰酸钾足迹，对于检测 TATA 盒结合蛋白（TBP）-TATA 相互作用已经特别有用（Giardina et al. 1992；Saunders et al. 2006)。

对于解链的 DNA 而言，高锰酸钾足迹最初用于确立在转录诱导之前真核细胞中的可诱导基因在其转录起始位点的下游拥有暂停的 RNA 聚合酶分子（见第 1 章）（Giardina et al. 1992；Lis and Wu 1993)。当聚合酶暂停在转录起始位点下游几十个碱基对处时，聚合酶分子附近的 DNA 模板保持解链状态。因此，未诱导细胞中对高锰酸钾修饰的高度敏感性可为暂停聚合酶分子存在的假设提供支持。

核小体存在和定位的 Southern 印迹分析

上述技术可以为转录因子和染色质蛋白与特定基因座的结合，或全基因组视角的蛋白质-DNA 相互作用提供有价值的信息。基本的 ChIP 技术对于监测特定位点或整个基因组上组蛋白的修饰也是有用的。然而，为了进行更先进的染色体结构研究，需要考虑其他的策略。这一节描述用于监测在特定基因组位置上核小体存在和定位的经典的微球菌核酸酶（MNase）-Southern 印迹方法。随后的一节讨论用于监测特定位置和全基因组水平上核小体密度和定位的高分辨率方法。然后，讨论基于核酸酶可及性监测体内核小体重塑的策略（在第 13 章中讨论用于核小体定位和重塑研究的体外方法)。

在目的控制区开始染色质结构的分析时，一个合理的出发点是确定该内源控制区是否被组装进核小体或缺乏核小体。第二个目标是确定核小体在细胞群体中是否定位在一致的位置。也就是说，在细胞群体中核小体在大多数等位基因上是相似地进行定位吗？在某些控制区上，据认为一致的核小体定位对于恰当的转录调控是关键性的（如见 Lomvardas and Thanos 2001)。虽然未彻底详细地弄清楚核小体定位的决定因素，但已经向确定那些促进核小体定位的特定 DNA 序列和蛋白质这一目标取得了大踏步的迈进（Pazin et al. 1997；Travers and Drew 1997；Vermaak and Wolffe 1998；Thástrom et

al. 1999；Ioshikhes et al. 2006；Segal et al. 2006；Rando and Ahmad 2007；Mavrich et al. 2008；以及其中的参考文献）。

用于监测核小体存在及核小体定位的大部分策略都利用了 MNase 的独特性质，该核酸酶优先在核小体连接区域或无核小体的 DNA 区域引入平末端双链断裂（重要的是要意识到 MNase 切刻核小体内的 DNA，限于在连接区和无核小体区域仅产生平末端、双链断裂）。如果在细胞核中加入极限浓度的 MNase，大多数基因组 DNA 将被消化成片段，这些片段在大小上对应于核小体核心区加连接区域长度（约 200 bp）的倍数。DNA 纯化后，可以通过琼脂糖凝胶电泳及随后的溴化乙锭染色来检测这些核小体梯带（图 9.5）。如果在细胞核中加入更高浓度的 MNase，较大比例的 DNA 将被消化成与大小上对应于核心单核小体（147 bp）的片段。

用于确定一个目的 DNA 片段是否被组装进核小体的经典方法涉及 MNase 消化 DNA 的 Southern 印迹分析，其中，Southern 印迹分析中使用含有目的 DNA 片段的放射性探针。如果探针与对应于核小体梯带的条带中的一条带杂交（见图 9.5，泳道 2），则目的片段可能被组装进核小体。如果观察到的不是一条条带而是放射性拖尾，或者如果没有明显的杂交，则该区域中可能没有核小体。

尽管此程序相对简单，但还是有两点重要的告诫。第一个告诫，由于凝胶上的大多数 DNA 是以 MNase 抗性核小体的形式存在，所以务必要论证探针是特异性地与目的基因组片段杂交，而不是非特异性的与大量 DNA 杂交。证实杂交特异性的一个方法是，在实验中包括一个来自不同物种细胞的 MNase 处理的 DNA 样品；当用溴化乙锭染色时，此 DNA 样品应该显示出相同的核小体梯带，但是不应与放射性标记的探针杂交（假定探针来源于种间差别足够大的 DNA 片段）。另一种方法是遵照专题 9.2 中描述的程序，该程序用于确定核小体是否定位。该方法与前面描述方法略有不同，但对非特异性杂交造成的人造品更不敏感。然而，只有当目的区域中的核小体确实定位于恒定位置时，此方法才会产生有用的信息。

第二个告诫是 MNase 抗性不会提供一个片段是核小体 DNA 的决定性证据。已经观察到近似正确大小的 MNase 抗性片段，据认为这些片段来自序列特异性 DNA 结合蛋白的紧密包装阵列（如见 Verdin et al. 1993）。因此，需要其他实验来支持 MNase 抗性片段的确是核小体 DNA 的假说。例如，如果假定核小体是精确定位的，并且其他核小体以与典型的串珠阵列相一致的间隔位于它的两侧，则 MNase 抗性片段可能确实代表核小体。基因激活时核小体重塑事件的证据也可以支持某片段是核小体 DNA 的假说，因为如果该区域缺乏核小体，这个用于监测重塑（如限制性内切核酸酶可及性分析，见下文）的分析方法通常会产生阴性结果。

为了确定与特定 DNA 片段相结合的核小体是否恒定定位，可以将 MNase-Southern印迹方法进行修改。该修改就是在 Southern 印迹分析之前用限制性内切核酸酶进一步消化已被 MNase 消化的基因组 DNA。专题 9.2 中对该方法作了更详细的描述。该分析的最大局限性是核小体的边界仅能被确定在大约 40 bp 以内，这归因于琼脂糖凝胶的分辨率及连接区中切割位点的可变性。因此，该技术的低分辨率可限制确定定位核小体相对于重要控制元件位置的精确性。

专题 9.2 通过 MNase-Southern 印迹分析进行核小体定位的低分辨率分析

MNase-Southern 印迹技术可以帮助确立 DNA 区域是否含有定位核小体的典型阵列。首先通过使用非离子型去污剂（如 NP-40 或 Triton X-100 等）裂解细胞膜，从合适细胞中制备细胞核（图 9.5）。通过离心分离细胞核后，等分细胞核，然后用不同浓度的 MNase

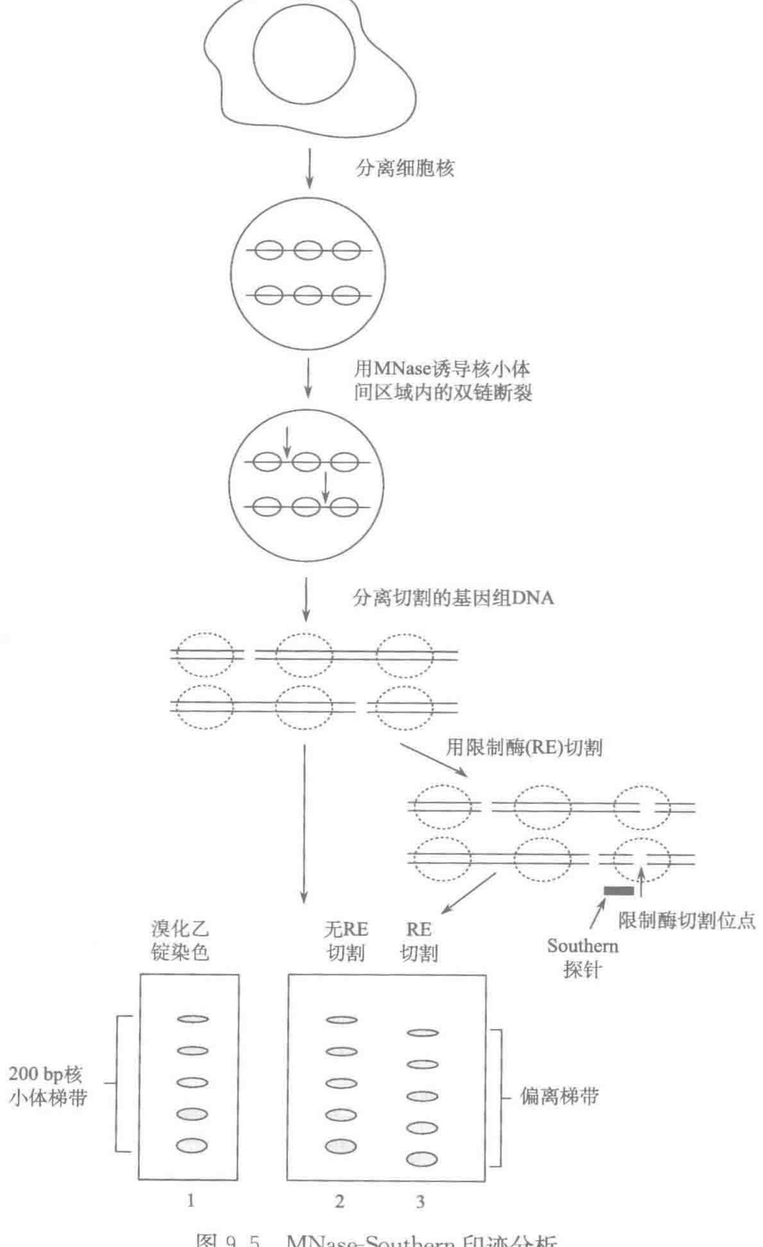


图 9.5 MNase-Southern 印迹分析。

处理每份细胞核。使用含有 SDS、EDTA 和蛋白酶 K 的溶液终止 MNase 的消化，然后纯化基因组 DNA，并用合适的限制性内切核酸酶切割，最后进行Southern印迹分析。

为了选择合适的限制性内切核酸酶，必须知道基因座的限制性内切核酸酶酶切图谱。理想的酶是那些在目的 DNA 区域附近切割的酶。之后需要一个放射性标记探针，该探针仅与限制性内切核酸酶酶切位点一侧的 DNA 序列杂交，并且杂交位置尽可能靠近限制性内切核酸酶酶切位点。紧邻切割位点的探针的杂交被称为“间接末端标记”，因为该结果与那些用^{32}P 直接标记被切割的 DNA 分子所获得的结果相似。

在缺乏限制性内切核酸酶切割的情况下，如果该区域组装成核小体，那么一旦探针杂交，MNase 消化的 DNA 就会显示出 200 bp 间隔的典型核小体梯带（见图 9.5，泳道 2）。如果核小体定位一致，用限制性内切核酸酶切割的样品就会产生一条新的梯状条带，该条带偏离标准梯状条带，其偏移的量取决于限制性内切核酸酶酶切位点相对于核小体的位置（见图 9.5，泳道 3）。例如，如果限制性内切核酸酶酶切位点位于核小体的中心，则新的产物应该为 100 bp、300 bp、500 bp 等，它们将会迁移到标准核小体产物 200 bp、400 bp 和 600 bp 之间的 1/2 处。如果限制性内切核酸酶酶切位点位于核小体间区，则限制性内切核酸酶酶切产物和标准核小体梯带可能难以区分，并且将有必要使用在核小体内部进行切割的不同限制性内切核酸酶重复该实验。如果这个区域不含有标准的串珠状核小体阵列，该结果会变得更加复杂。如果核小体定位不一致，限制性内切核酸酶切割应导致 Southern 印迹图像拖尾的增加。如果多个限制性内切核酸酶不能产生提示定位的条带，该区域就可能没有核小体，或者可能含有定位不一致的核小体。

监测核小体存在和定位的 LM-PCR、PCR 及 ChIP 策略

用于检测核小体存在和定位的方法，除了传统的 Southern 印迹方法外，还可以使用 LM-PCR、PCR 和基于 ChIP 的方法。这些方法和 Southern 印迹相比通常产生关于核小体定位的更高分辨率的信息，并且其中某些方法可以在全基因组水平上进行。和 Southern 印迹一样，这些方法很大程度上依赖于使用 MNase 切割位于核小体位置之间和无核小体区的染色质。

这些策略中的有些策略始于用 MNase 将染色质消化成单核小体，然后使用凝胶电泳（如见 Lomvardas and Thanos 2001；Sekinger et al. 2005；Schones et al. 2008）或者梯度离心（O’ Neill and Turner 2003）纯化单核小体。作为一种替代方法，可以用针对未修饰组蛋白的抗体进行 ChIP 实验，以分离 MNase 抗性核小体（Mavrich et al. 2008）。在缺乏组蛋白与 DNA 交联的情况下，可执行两种程序，或者先进行甲醛交联，再执行两种程序；接着进行单核小体纯化或免疫沉淀反应，最后进行解交联（O’ Neill and Turner 2003；Thorne et al. 2004）。在不进行交联的情况下，核小体足够稳定，以致可进行单核小体纯化或者免疫沉淀；然而，在 MNase 消化期间，核小体可能会发生重新定位，因此，已经产生了关于选择性核酸酶消化的担忧（O’ Neill and

Turner 2003)。用于分离核小体 DNA 的 ChIP 方法的主要优点是比纯化方法速度更快，并且技术上更简单；纯化方法有时导致纯化 DNA 的低产量。然而，单核小体的纯化可能产生高质量的数据，因为 ChIP 策略更有可能导致高背景信号，原因是 MNase 并未有效消化较大 DNA 片段的沉淀。值得注意的是，虽然 MNase 通常用来研究核小体的密度和定位，但是也有可能使用一个基本的 ChIP 程序，在该程序中，通过超声剪切甲醛交联的染色质，然后使用针对未修饰组蛋白的抗体进行免疫沉淀反应（Heintzman et al. 2007）。该程序可以提供特定基因组位置上关于核小体密度的信息，但是它不能提供关于核小体位置的高分辨率信息。

可以通过多种程序中的任何一种程序分析由单核小体纯化或 ChIP 分离的 MNase 抗性 DNA 片段，其中包括使用跨越目的区域的一系列引物对的实时 PCR 扫描（Sekinger et al. 2005）、LM-PCR（方案 9.2）（McPherson et al. 1993；Weinmann et al. 1999）、引物延伸（Lomvardas and Thanos 2001）、扩增引物延伸（Sekinger et al. 2005）、叠连阵列杂交（Heintzman et al. 2007）和大规模平行测序（Mavrich et al. 2008；Schones et al. 2008）。

这些用于 DNA 分析的不同方法将会产生分辨率不同的结果。理论上讲，PCR 扫描方法产生分辨率最低的数据，对于这种方法，要使用跨越目的区域的大量 PCR 引物对。在 Sekinger 等（2005）的研究中，每个引物对扩增大约 100 bp 的片段，每一个片段位于距邻近片段大约 30 bp 处。图 9.6 显示了使用该方法对良好表征的酿酒酵母（*Saccharomyces cerevisiae*）*Pho*5 启动子进行分析所获得的数据。在这个研究中，在分析之前需要通过凝胶电泳纯化单核小体 DNA。图底部显示了用于分析的 PCR 扩增片段，之前被定位的核小体用椭圆形表示，使用 PCR 扫描分析获得的数据用条形图显示（作为 MNase 消化基因组 DNA 的相对扩增）。扩增片段的 30 bp 重叠大大提高了分析的分辨率，但是这个分辨率仍然低于使用其他检测方法所获得的分辨率。

和 PCR 扫描方法相比，当在变性聚丙烯酰胺测序凝胶上分析，并以邻近的序列梯带作为分子质量标准物时，LM-PCR、引物延伸和扩增引物延伸分析有能力在只有少数几个碱基对的分辨率上定位 MNase 切割位点。图 9.7 显示了一个在小鼠 *Il12b* 启动子分析中获得的 MNase-LM-PCR 结果的实例。在该研究中（Weinmann et al. 1999），用 MNase 消化基因组 DNA，在没有进行单核小体纯化和免疫沉淀的情况下，对其直接进行 LM-PCR（专题 9.2）分析，结果发现高敏感性切割位点的两个明显的紧密包装簇，这些包装簇可能对应于核小体之间的区域，在这些区域中 MNase 已经引入了双链断裂，这些位点间的 135 bp 保护区域大约是核小体核心的大小。通过确定放射性标记片段的大小［通过与在邻近泳道上电泳的分子质量标准物比较而确定（未显示）］，以及对放射性标记引物杂交的基因组位置的了解，可以以相当大的精确性确定这些假定核小体间切割的位置（在这种情况下，相对于转录起始位点的核苷酸－350～－485）。引物延伸方法和扩增引物延伸方法应提供相似的分辨率。

其余的两种检测方法，即叠连阵列杂交和大规模平行测序，可提供全基因组水平上核小体定位的信息（如见 Heintzman et al. 2007；Mavrich et al. 2008；Schones et al. 2008）。大规模平行测序的使用可提供分辨率极高的 MNase 切割位点信息，甚至超过了使用叠连阵列所获得的分辨率信息，这是因为 DNA 序列可以揭示 MNase 抗性片

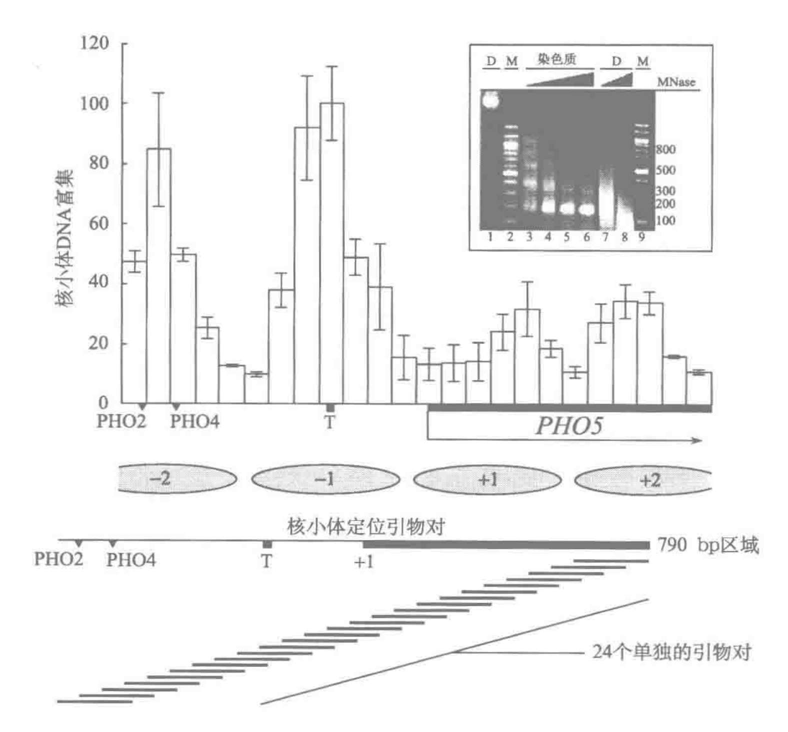


图 9.6　酿酒酵母（*S. cerevisiae*）*Pho5* 启动子上核小体定位的 PCR 扫描分析（经允许转载自 Sekinger et al. 2005，© Elsevier）。

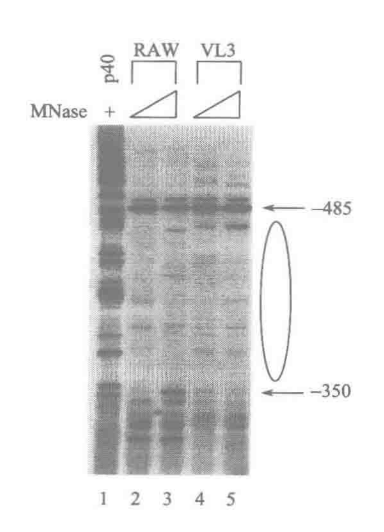


图 9.7　小鼠 *Il12b* 启动子上核小体定位的 MNase-LM-PCR 分析（经允许转载自 Weinmann et al. 1999，© Elsevier）。

段的准确序列。然而，这两种方法都已经相当成功地用于监测核小体的密度。

重要的是要注意，即使使用高分辨率的基于 MNase 的方法，对于 MNase 切割位点精确位置的认识并不能提供对核小体位置的绝对认识。首先，MNase 切割位点的位置部分由 DNA 序列决定。MNase 不会在核小体之间的区域内随机切割，而是具有固有的序列偏好性。因此，虽然大多数对于 MNase 的切割高度敏感的位点位于核小体之间的区域内，但是不能确定核小体两侧的准确边界。其次，核小体定位可能很灵活，并因此在生长细胞内可能不会严格地限制在精确的位置上（Fragoso et al. 1995）。

除了 MNase-LM-PCR 方法，DNase I 基因组足迹法可有助于了解核小体定位。DNase I

优先切刻背对组蛋白八聚体的螺旋 DNA。因此，当螺旋 DNA 缠绕核小体时，可以在大约 10 bp 的间隔上观察到稍微高敏感性的切割。如果观察到这些切割，就可以提供关于核小体旋转相位的信息（见图 13.3）。

用于分析核小体重塑的体内方法的概述

正如第 1 章中所描述的，核小体重塑在许多真核基因的调控中发挥着重要作用。由于这个原因，研究染色质与基因转录之间关系的研究者，常常对监测基因从转录失活状态转到转录激活状态（或反之亦然）时染色质结构中的物理变化感兴趣。除了使用 ChIP 研究染色质蛋白与组蛋白翻译后修饰之间的关系外，用于监测体内染色质结构的最普遍的方法涉及一项 DNA 区域对内切核酸酶可及性的分析，如 DNase I、MNase 或序列特异的限制性内切核酸酶等内切核酸酶。核酸酶可及性实验通常以分离的细胞核或者通透化的细胞进行；在通透化细胞中，由于不存在质膜障碍，核酸酶能够接触染色质[已经发展了在活细胞内监测染色质可及性的策略，但使用相对较少（见 Gottschling 1992）]。

内切核酸酶在特定基因组区域上切割的效率通常与染色质状态有关。对于压缩染色质以相对较低的切割效率切割，而对于更加开放的染色质则以高效率切割，尤其对于那些缺乏核小体的染色质，或者被三磷酸腺苷（ATP）依赖性核小体重塑复合物重塑了核小体的染色质，则以更高的效率切割。生物化学研究已经令人信服地证明，由 ATP-依赖性重塑复合物执行的核小体重塑，以及核小体的驱逐，可以大幅提高对于内切核酸酶切割的敏感性（见第 13 章）。遗憾的是，当发现基因组区域对于切割相对可及或者不可及时，在没有更高级实验的情况下，不能确定精确的染色质结构和负责其特定可及性水平的染色质重塑复合物。尽管如此，基因激活或沉默时特定基因组区域可及性变化的相关证据，可为阐明潜在的结构性变化、造成这些变化的蛋白质复合物及其与基因转录调控的相关性提供有价值的出发点。

DNase I 高敏感性监测核小体重塑

第 5 章中介绍了 DNase I 敏感性和 DNase I 高敏感性两种核酸酶可及性分析方法。发展 DNase I 敏感性方法的最初目的是用于监测基因座总体染色质结构的变化，它是 DNase I 高敏感性分析方法的前身。第 5 章中对 DNase I 高敏感性分析方法进行了描述，因为该方法对于鉴定真核基因远程控制区是有用的，这些真核基因往往无核小体或者与重塑核小体相结合。当基因被激活或沉默时，DNase I 高敏感性的变化提供了伴随激活或沉默的染色质结构改变的证据。

第 5 章所描述的传统 DNase I 高敏感性分析对于鉴定远程控制区和确定控制区的一般可及性是否受到调控是有用的。然而，通常不推荐将该分析用于详细分析核小体组建和重塑，因为这种分析相对来说不灵敏，并且有时无法检测到染色质可及性的增加，而

使用下面所描述的限制性内切核酸酶可及性分析则可以检测到。另外，利用基因组 Southern 印迹来检测切割产物的传统 DNase I 高敏感性分析，仅提供 DNase I 切割位点低分辨率的信息。大多数研究核小体重塑的研究人员对在较高分辨率上鉴定经历重塑的特定核小体和 DNA 区域感兴趣。

造成传统 DNase I 高敏感性分析方法分辨率低的原因是，DNase I 通常在高敏感性控制区内的多个位置上进行切割。为了使用 Southern 印迹提供 DNase I 切割位点的高分辨率信息，必须对 DNase I 切割的基因组 DNA 进行纯化，然后用识别高敏感区附近已知位点的限制性内切核酸酶进行消化（见第 5 章）。理论上讲，之后可以用合适的分子质量标准物进行 Southern 印迹，以确定从已知限制性内切核酸酶酶切位点到 DNase I 切割位点的 DNA 片段的准确大小。不幸的是，这种高分辨率策略在实践中通常是不成功的，因为只有小比例的基因组 DNA 分子在高敏感区内若干 DNase I 切割位点的每一个位点上被切割（可能全部 DNA 分子中 2%～5%的 DNA 分子在每个 DNase I 切割位点处被切割），由于基因组 Southern 印迹技术固有的不灵敏性，使用放射性标记的 Southern 印迹探针检测单个 DNA 片段就会变得非常困难。

由于这个局限性，DNase I 高敏感性分析方法通常使用离广泛的 DNase I 切割高敏感性区域至少 1 kb 距离外切割的限制性内切核酸酶，之后切割产物在低分辨率的琼脂糖凝胶上电泳分析。通过这个策略，DNase I 在多个成簇位点上切割产生的 DNA 片段合并成一个易检测的条带。然而，对于精确的 DNase I 切割位点，该结果只提供有限的信息。

虽然上述技术上的局限性使得利用传统 DNase I 高敏感性分析方法在高分辨率上研究核小体重塑变得困难，但是 DNase I 切割可以和第 5 章及本章中所描述的 LM-PCR 或者其他基于 PCR 的检测方法联用，也可以在全基因组规模上进行 DNase I 高敏感性分析（见第 5 章）。然而，由于下文所强调的原因，大多数研究人员仍发现使用在目的 DNA 区域内切割的特定限制性内切核酸酶监测核小体重塑更为实际。

用于监测核小体重塑的 MNase 方法

用于研究核小体定位的 MNase-Southern 印迹分析、MNase-LM-PCR 分析及其他基于 MNase 的分析（见上文），也可以用于监测核小体的重塑。如上所述，组装成稳定核小体的 DNA 通常对 MNase 的双链切割具有抗性，但是在核小体已经被驱走的 DNA 区域或者含有重塑核小体的 DNA 区域中，会表现出增强的 MNase 可及性。为了达到这些目的，在使用 MNase 进行分析时，可将基因表达的细胞与基因不表达的细胞相比较。表达细胞中对 MNase 切割敏感性的增加，可提供核小体重塑的证据。当重塑事件导致核小体滑动到不同位置时，MNase 分析方法依然有用（Lomvardas and Thanos 2001）。在这种情况下，核小体定位实验将证明重塑前后核小体定位在不同的位置。

虽然 MNase 分析方法在某些情况下可能有用，但是它们具有与上述 DNase I 分析方法相似的技术上的局限性。例如，如果一个基因在一种细胞系内诱导无效（这是十分常见的），则在一小部分细胞内增加的切割可能被大多数细胞中变化的缺乏所掩盖，尤其是当这个问题和由重塑核小体内大量潜在 MNase 切割位点造成的该分析方法有限的灵敏度相结合时。为了说明这些局限性，图 9.8 显示了重塑前后对 *Il12b* 启动子分析所得到的 MNase-LM-PCR 结果（A. S. Weinmann and S. T. S.，未发表）。在这个例子中，基因诱导时，确定的核小体内的 MNase 切割仅略微提高，从而产生暗示核小体发生重塑的结果，但是这些结果不如下面所描述的限制性内切核酸酶可及性结果那样令人信服。值得注意的是，基因诱导后最强的 MNase 切割位点仍然保持不变，并且没有出现其他显著的位点，这与核小体在重塑期间不会滑动到不同位置的假设是一致的。这种适度影响的一个可能的原因是，细胞在受刺激后，*Il12b* 基因仅在这一克隆细胞系的 25% 的细胞中表达。对此起作用的第二个因素是，重塑后核小体内的每一个 MNase 切割事件仅发生在该群体内一小部分的 DNA 分子上，从而导致弱信号，即使当使用敏感的 LM-PCR 检测方法时，也是如此。

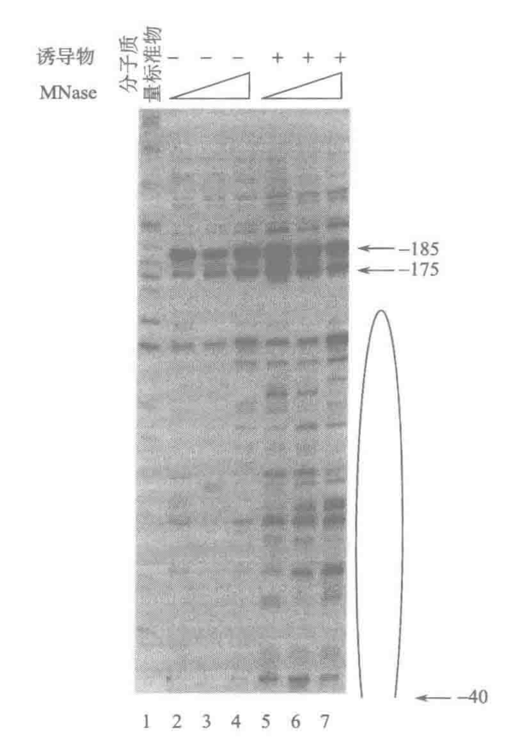


图 9.8　LPS 转录诱导前后小鼠 *Il12b* 启动子区的 MNase 切割模式。诱导前，通过在核苷酸 −175 和 −185（泳道 2～4）处的高敏感性切割位点，可揭示一个假定核小体间区。诱导后，从这些位点的下游区域，在 −40 和 −175 之间，对 MNase 的切割变得更加敏感，这与核小体重塑相一致（泳道 5～7）（A. Weinmann and S. T. Smale，未发表）。

限制性内切核酸酶可及性分析

限制性内切核酸酶可及性分析是多种分析方法的集合，这些分析常常可克服用于监测核小体重塑的 DNase I 和 MNase 方法的局限性，但这些分析也有自身的局限性。这些分析方法可监测限制酶可接近和切割分离的细胞核或者通透化细胞中的目的 DNA 区域内已知的识别位点的效率。相对于 DNase I 和 MNase 分析方法，限制性内切核酸酶可及性分析的一个主要优点是，其结果提供特定位点上的可及性信息，这大大提高了检测的灵敏度。而且，因为可以从 DNA 序列中确定精确的切割位置，所以很少遇到切割位置作图的困难。最后，由于切割仅发生在一个位点而不是在控制区内的若干分散位点上，更容易量化结果。

然而，该分析方法的一个缺点是被监测的特定位点可能定位在核小体内的一个位置上，这提供了关于伴随基因激活或沉默的染色质结构变化的误导性信息。例如，该位点可能与转录因子结合位点相一致，当该因子结合时导致切割减少。此外，核小体重塑和基因激活之前，对切割来说，限制性位点可能相对易于接近，因为限制性位点可能定位于核小体连接区，或由于其在核心核小体内的旋转定位，限制性位点可能暴露于酶。然而，重塑后，由于核小体的滑动可能导致切割不太易于接近限制性位点。因此，重塑会使该位点的可及性降低，这与所预期的可及性增加相反，从而可能导致研究者误解该结果。由于这个局限性，强烈推荐尽力监测目的区域内多个不同的限制性位点，在上文所述的例子中，这将有可能揭示核小体的滑动事件。遗憾的是，限制性内切核酸酶可及性分析还有第二个局限性，那就是在许多目的 DNA 区域内发现合适的限制性位点并不容易。

限制性内切核酸酶可及性分析-Southern 印迹

当进行限制性内切核酸酶可及性分析时，最可靠的检测方法是 Southern 印迹。首先用限制酶处理细胞核或通透化细胞，之后纯化基因组 DNA，再用另外一个或两个限制酶进行体外切割，目的是挑选能在细胞核切割位点的每一侧产生切割的限制酶。当使用与细胞核切割位点一侧序列互补的探针进行 Southern 印迹分析时，将检测到群体中所有的 DNA 分子。一部分 DNA 分子已经在细胞核内被切割，而剩余的分子将只能在体外进行切割，前者的片段将比后者的小。因为所有 DNA 分子被检测，有可能相当自信地计算出每个样品细胞核内被切割的 DNA 分子的百分比。因此，与下面描述的替代策略不同，该策略含有理想的内对照，可在很大程度上消除对特定实验中许多样品处理中变化性的担忧。另外一个重要好处是，由此产生的细胞核切割比例总是不受 PCR 扩增的影响，从而允许在特定实验中对许多样品相互间进行有意义的比较。

Southern 印迹检测方法的主要局限性是需要使用大多数情况下识别 6 bp 序列（6 bp 切点）的限制酶，这在许多控制区内可能是很难找到的。通常不能使用识别 4 bp 序列（4 bp 切点）的酶，因为这种酶产生小尺寸 DNA 片段，这些 DNA 片段在 Southern 印迹上很难检测。

图 9.9 显示了对鼠 *Il12b* 启动子和增强子分析得到的限制性内切核酸酶可及性-Southern 印迹结果的例子（Ramirez-Carrozzi et al. 2006）。在该实验中，将 6 bp 切点的限制性内切核酸酶 *Spe*I 用于细胞核切割。该酶的两个识别位点定位于启动子中（在 −82和 −107 处，它们产生在 Southern 印迹不能分解的条带），还有一个另外的位点在增强子中（相对于转录起始位点，在 −10 kb 处）。巨噬细胞未刺激或用脂多糖刺激不同时间（在蛋白质合成抑制剂放线菌酮的存在下再刺激 120 min，以确定重塑是否依赖于蛋白质的合成）后，制备细胞核并用 *Spe*I 孵育。纯化基因组 DNA，之后用已知识别启动子和增强子两侧位点的其他限制酶完全切割（未显示）。之后使用和 *Spe*I 切割位点邻近部位杂交的启动子或增强子特异性放射性标记探针进行 Southern 印迹分析。上部条带对应于对纯化的 DNA 进行切割而产生的 DNA 片段。较低的条带显示了当细胞核内的等位基因也被 *Spe*I 切割时产生的 DNA 片段。磷屏成像分析允许定量细胞核内被 SpeI 切割的等位基因所占的百分比。由该定量分析提供的一个认识是，基因活化后，

对于 *Spe*I 切割来说，*Il12b* 增强子比启动子更易接近。在较早的 LM-PCR 实验中，这种在可及性上的差异并不明显（见下文）；由于 PCR 扩增，LM-PCR 实验会产生不太定量的结果（后续的研究表明增强子可及性更好，因为在基因活化过程中核小体从该位置被驱逐，而在启动子中核小体不被逐出）。

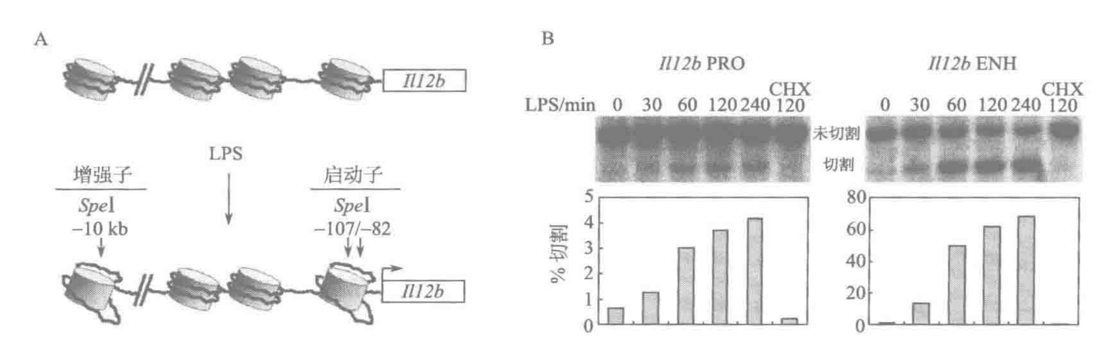


图 9.9 小鼠 *Il12b* 启动子和增强子限制性内切核酸酶可及性-Southern 印迹分析（经允许转载自 Ramirez-Carrozzi et al. 2006）。

限制性内切核酸酶可及性-LM-PCR 分析

作为 Southern 印迹检测方法的替代方法，LM-PCR 可用于限制性内切核酸酶可及性分析（专题 9.3）。正如 Southern 印迹方法，在基因组 DNA 作为内对照纯化后，用第二个限制酶体外切割基因组 DNA 也是有益的；基因组 DNA 内对照使特定 LM-PCR 实验中多个样品之间的直接比较成为可能。然而，在这种情况下，需要体外切割仅发生在细胞核切割位点的一侧。

图 9.11 显示了以 *Il12b* 启动子进行的限制酶 LM-PCR 分析的一个例子。结果显示，受刺激的巨噬细胞核内（泳道 8～13）三种不同限制性内切核酸酶 *Spe*I、*Mse*I 和 *Dde*I 的切割大大地高于未受刺激的巨噬细胞核内（泳道 1～7）三种酶的切割。在这些实验中，纯化的 DNA 用 *Aat*II（泳道 2、3、8 和 9）或 *Pst*I（泳道 4～7 和 10～13）切割，以达到归一化目的。例如，将泳道 2 和泳道 3（二重反应）与泳道 8 和泳道 9 的比较，揭示出 *Spe*I 对 *Aat*II 的切割比率在受刺激样品中（泳道 8 和 9）比未刺激样品中（泳道 2 和 3）更高。

LM-PCR 分析的一个优点是它比 Southern 印迹分析灵敏，并且比 Southern 印迹更容易揭示弱切割事件。然而，LM-PCR 的一个主要缺点是大规模的 PCR 扩增使研究者不能确定细胞核内被切割分子的百分比，并且不能比较样品与样品之间的那些值。LM-PCR 分析的第二个缺点是该分析比 Southern 印迹方法更难并且更耗时。

实时 PCR 方法

可将另外两种基于 PCR 的方法用于限制性内切核酸酶可及性分析。一种方法最初命名为 CHART-PCR 分析（通过实时 PCR 的染色质可及性）（Rao et al. 2001），它利用位于限制性内切核酸酶位点两侧的 PCR 引物。和其他方法一样，用限制酶孵育分离的细胞核或者通透化的细胞，接着纯化 DNA。如果基因组 DNA 在目的位点被酶有效

切割，与 DNA 未被酶有效切割相比，对于使用侧翼引物的 PCR 扩增来说，则有更少的完整模板分子。可将实时 PCR 用于定量每个样品中仍未切割的 DNA 的量。

CHART-PCR 分析的主要优势在于它比 Southern 印迹或者 LM-PCR 分析更为简单。此外，像 Southern 印迹分析一样，它可以提供特定样品中已被限制酶切割的基因组 DNA 分子的百分比的定量信息。然而，和 Southern 印迹方法不同的是，CHART-PCR 分析缺乏内对照，因为被切割分子的百分比必须通过和单独的模拟样品（未经切割的基因组 DNA）进行比较才能确定。通过使用实时 PCR，对从实验样品及模拟样品中获得的信号进行归一化，从而监测每个样品中的不同的基因组 DNA 区域，这些基因组 DNA 区域中不含有正被使用的限制酶的识别位点。

当控制适当时，该分析会非常有用，但主要是对一些特殊位点有用，包括核小体重塑后限制酶切割效率呈现巨大变化的位点，以及重塑状态下切割效率相对高的位点（大于 50%）。例如，在 Rao 等（2001）进行的 *Il2* 启动子的研究中，在 *Il2* 基因无活性的未刺激细胞中 0%的等位基因常常被切割，而在受刺激细胞中 50%～100%的等位基因被切割。这种可及性的戏剧性变化与后来的发现是一致的，即 *Il2* 的激活与核小体驱逐相一致（Chen et al. 2005）。然而，在我们的经验中，基因激活时在大部分控制区所观察到的可及性变化的幅度要小得多（如见 Ramirez-Carrozzi et al. 2006）。例如，在图 9.9 所示的限制性内切核酸酶可及性-Southern 印迹实验中，*Il12b* 启动子的可及性从大约 0.5%增加到 4%。在内部控制的 Southern 印迹分析方法中，这 8 倍的变化是显而易见的，但是在 CHART-PCR 分析中则检测不到。在 CHAR-PCR 分析中，从未刺激状态中的 0.5%的切割变到受刺激状态中的 4%的切割，就相当于 99.5%未切割等位基因变到 96%的未切割等位基因，这将导致在实时 PCR 分析方法中小于 0.1 个循环的差异，这种差异远低于标准实验的变异。因此，尽管 CHART-PCR 分析对于显示可及性戏剧性变化的控制区（如 *Il2* 启动子及 *Il12b* 增强子）的分析（见图 9.9）是有用的，但是对于我们已经检测到的大多数控制区是没有用的（V. Ramirez-Carrozzi, A. Nazarian, and ST. Smale，未发表）。

专题 9.3　限制性内切核酸酶可及性-LM-PCR 分析

限制性内切核酸酶可及性-LM-PCR 分析（图 9.10）是一种用于监测特定核小体上核小体重塑事件的方法。如上所述（见专题 9.2），该分析涉及通过去污剂裂解来制备完整的细胞核。之后，在典型的限制性内切核酸酶消化缓冲液中，用极限浓度的限制性内切核酸酶孵育细胞核。重要的是要进行相对短时间的消化，以便尽可能地降低其对核小体完整性的破坏。纯化酶切割的基因组 DNA，用另一种作为内标的限制酶彻底切割。该酶应该是一种相对靠近第一种酶切割的酶，以便在相同的 LM-PCR 分析方法中可监测这两种体外切割产物。继 LM-PCR 程序（见专题 9.1）之后，通过用总放射性（即核切割产物加上体外切割产物）除以细胞核内切割产物的放射性，将可及性表示为分数形式，这个比值可能未精确反映核内被切割的部分（因为 LM-PCR 技术不是线性的），但对从基因表达和基因不表达的细胞中获得的结果进行比较，可获得一定的信息。

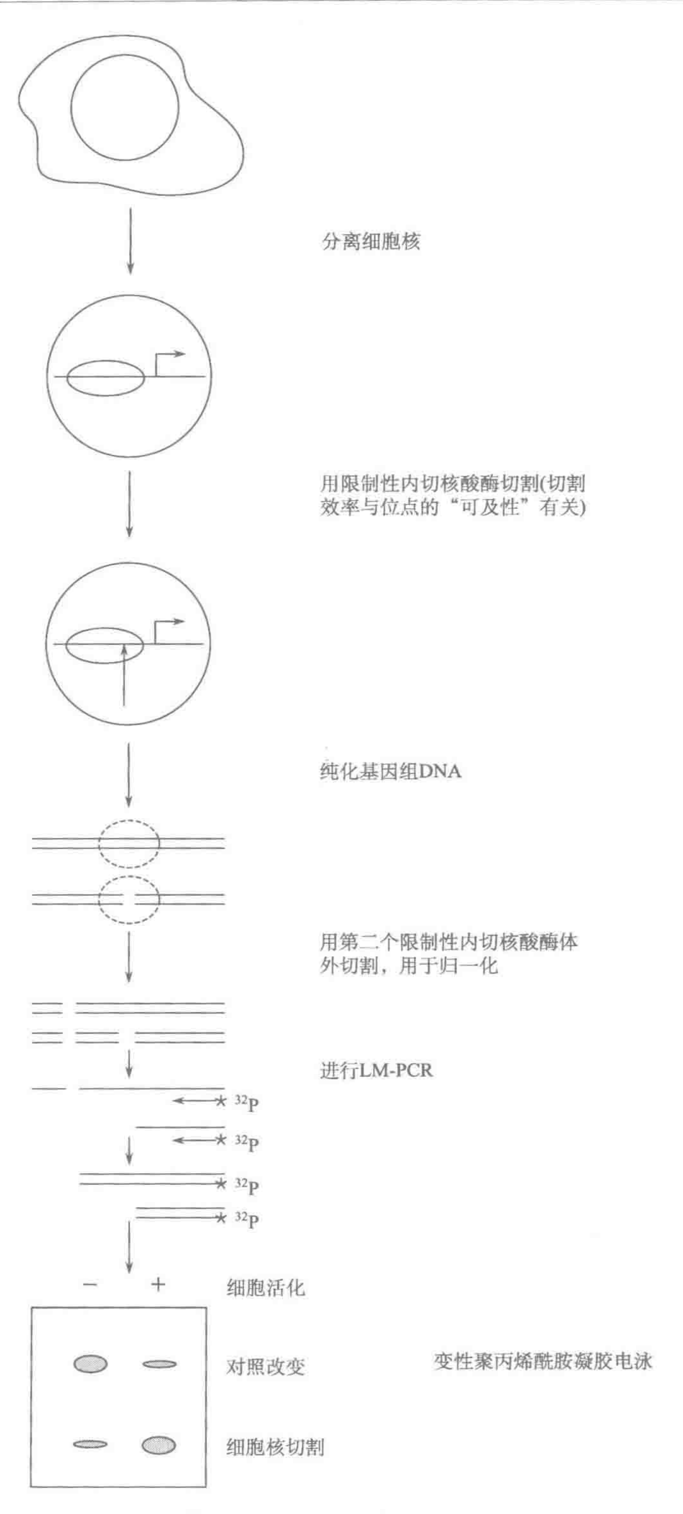


图 9.10　限制性内切核酸酶可及性-LM-PCR 分析。

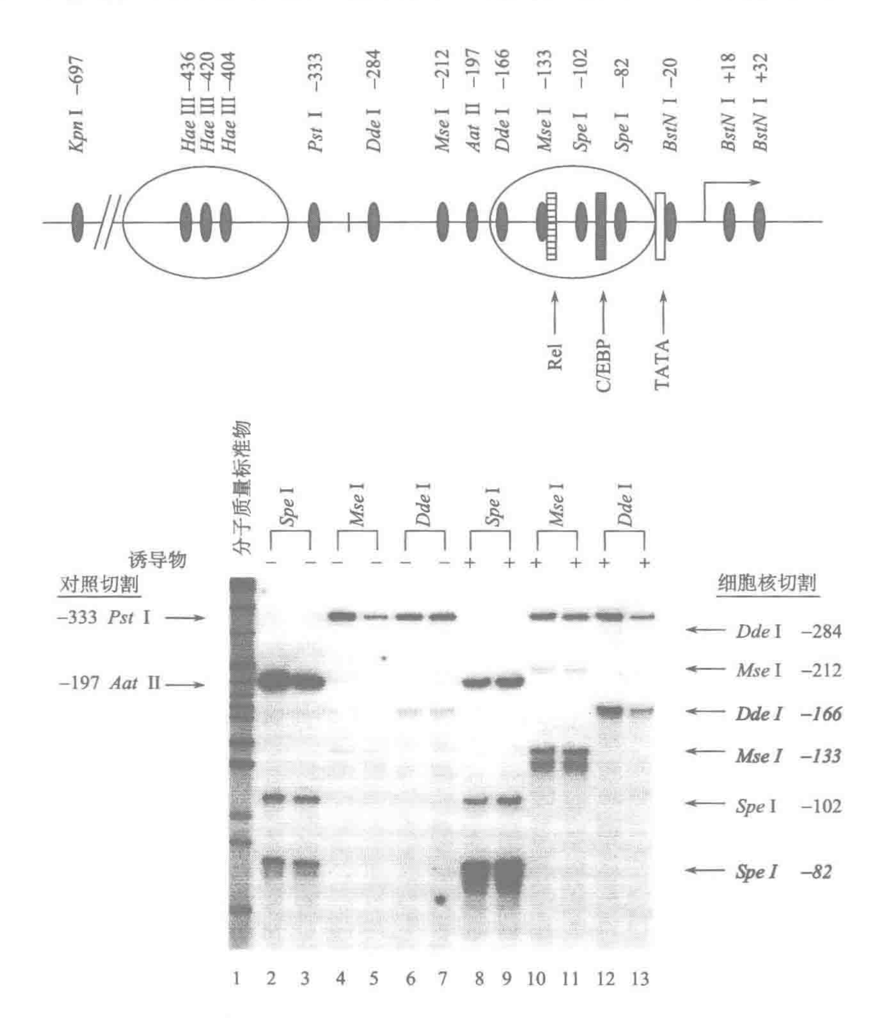


图 9.11 小鼠 *Il12b* 启动子限制性内切核酸酶可及性-LM-PCR 分析。结果显示，来自活化细胞的细胞核内的限制性内切核酸酶切割（使用 *Spe*I、*Mse*I 和 *Dde*I 酶；泳道 8～13）产生的 LM-PCR 产物，比使用相同酶对未活化细胞的细胞核进行切割后所观察到的 LM-PCR 产物更丰富（泳道 2～7），接着用体外对照进行归一化处理（*Pst*I -333 和 *Aat*II -197）（经允许转载自 Weinmann et al. 1999，© Elsevier）。

为了克服 CHART-PCR 分析的这种局限性，可以使用基于 PCR 的分析，该分析中细胞核切割的增加引起了 PCR 模板丰度的增加而非减少。该替代分析类似于 LM-PCR 分析，不同的是用实时 PCR 监测和定量 PCR 扩增，而不是使用依赖于凝胶电泳和磷屏成像分析（或放射自显影）的最后的放射性标记 PCR 步骤进行检测。首先用限制性内切核酸酶孵育细胞核，之后纯化 DNA 并将接头连接到切割末端。接着，进行实时 PCR，而非进行其他的 PCR 及用于 LM-PCR 的凝胶电泳。实时 PCR 中使用与目的基因组区域退火的正向引物和与接头退火的反向引物。

理论上讲，对于监测相对低的切割效率，该分析应比 CHART-PCR 分析更具吸引

力。例如，如果重塑之前有 0.5%的 DNA 分子被切割，重塑之后有 4%的 DNA 分子被切割，则可以检测到实时 PCR 产物丰度的 8 倍变化。然而，实际上当尝试使用该分析时，重塑前我们已经观察到的信号要比预期的高，这归因于相当大的背景（D. Bhatt and S. T. Smale，未发表），这些背景信号可能使该分析不如预期的那样理想，除非付出相当大的努力建立使背景最小化的实验条件。

CHART-PCR 分析和这个替代分析的一个主要优点是，它们可以很容易地与识别 4 bp 序列（4 bp 切点）的限制酶联用，因为切割产物的检测不依赖于两侧基因组区域中酶识别序列的缺乏。另外，这些分析可与 DNase I 或者 MNase 联用，从而提供非特异性内切核酸酶可及性相关的信息。然而，对于大多数具有恰当的限制性内切核酸酶识别序列的控制区，限制性内切核酸酶可及性-Southern 印迹分析仍然是首选分析，因为该分析能检测相对小的可及性变化，并且内对照的存在为该分析增强了信心。

染色质构象捕获

正如第 1、第 2 和第 5 章中所描述的，现在已提供了清楚的证据，表明单个基因的多个控制区之间可通过成环及形成活性染色质中心（ACH）而相互作用（Tolhuis et al. 2002；Spilianakis and Flavell 2004；de Laat and Grosveld 2007；Simonis et al. 2007；de Laat et al. 2008)，还检测到位于不同染色体上的基因相互之间靠得很近，说明染色体间的相互作用对转录调控可做出重要贡献（Spilianakis et al. 2005；de Laat et al. 2008)。支持染色体内及染色体间相互作用的证据主要来自被称为染色质构象捕获（3C）的技术。

3C 技术是由 Nancy Kleckner 及其同事（Dekker et al . 2002）发展的，作为一种通用方法，可将其用于监测不同基因组区域之间互作的频率。在原创研究中，该技术用于研究酵母（*S. cerevisiae*）细胞周期的减数分裂及 G_1 期染色体组织的变化。之后不久，Grosveld、deLaat 及其同事对该技术进行了修改，以便在哺乳动物细胞中使用，并证明了小鼠 β-珠蛋白基因座的不同控制区间的相互作用（ Tolhuis et al. 2002 ）。

简而言之，3C 技术始于利用甲醛进行活细胞内蛋白质与 DNA 及其他蛋白质之间的交联（图 9.12)。交联的 DNA 之后用限制性内切核酸酶彻底消化。在稀溶液中，加入 DNA 连接酶和 ATP，将 DNA 的游离端与之靠近的其他游离端相连接。由于连接在稀溶液中进行，如果两个 DNA 分子彼此交联（通过蛋白质-DNA 和蛋白质-蛋白质交联)，则连接一般将以高效率发生，因为在细胞核中它们靠得很近。之后，目的 DNA 区域间的相互作用用该区域特异的引物进行 PCR 监测。

在 3C 技术上又发展了若干较新的方法，包括 4C、5C 和 ChIP-环（图 9.12）（见 Simonis et al. 2007，以及其中的文献；Göndör et al. 2008)。4C 技术使用反向 PCR 扩增大量与目的片段连接的基因组片段。然后，PCR 产物可以与微阵列进行杂交或者通过大规模平行测序进行分析，以获得与目的位点相互作用的基因组区域的全基因组视图。5C 技术发展为一种策略，用于增加进行单个基因座或者扩展基因组区域详细表征的可行性。在传统 3C 分析中，一般设计与跨越基因座的很多区域退火的 PCR 引物，之后进行单个的 PCR，以检查每个位点与固定位点之间的连接。在 5C 程序中，用多重

PCR 同时扩增所有 PCR 产物，之后进行微阵列杂交或大规模平行测序。最后，通过在 DNA 连接之前免疫沉淀交联的、片段化的 DNA，用 ChIP-loop 程序分析发生在目的蛋白特异性结合的 DNA 区域间的相互作用。Simonis 等（2007）很好地总结了每种方法的使用范围及其局限性，并引用了详细的方案。

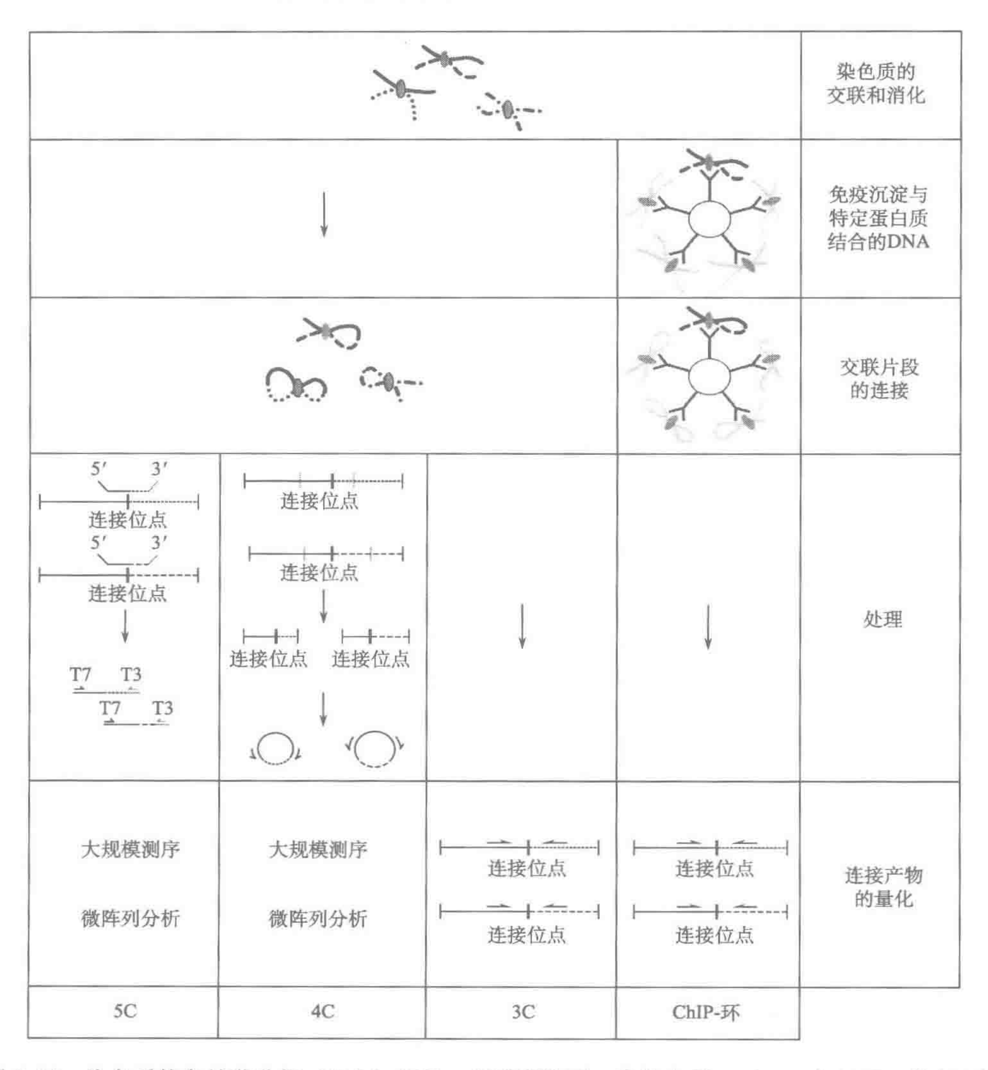


图 9.12 染色质构象捕获分析（经 MacMillan 出版商许可，重绘自 Simonis et al. 2007，© 2007）。

3C 方法现在已经非常成功地用于研究在许多基因座上的相互作用。使用在特定控制区内含有靶突变的细胞或者含有编码重要调控因子基因缺失的细胞，已经开始对负责 ACH 形成的 DNA 序列和转录因子有了深入的了解（Drissen et al. 2004；Patrinos et al. 2004）。3C 结果也已经开始揭示细胞发育期间染色质架构的动态变化（Jing et al. 2008）。未来的主要挑战将是阐明特定 DNA 区域的物理接近促成精确转录调控的机制。

DNA 甲基化

正如组蛋白翻译后修饰那样，DNA 甲基化被认为在基因表达的表观遗传调控方面发挥关键作用（第 1 章）（Klose and Bird 2006；Suzuki and Bird 2008）。在哺乳动物基因组中，DNA 甲基化主要限制在 CpG 二核苷酸的胞嘧啶中。然而，在一些真核生物[如拟南芥（*Arabidopsis thaliana*）] 中，其他 DNA 模体通常也被甲基化（Chan et al. 2005），但是很多其他真核生物（如酵母、果蝇和蠕虫）则缺少显著的 DNA 甲基化。正如第 1 章中所描述的，哺乳动物细胞内的 DNA 甲基化是由 DNA 甲基化转移酶 DNMT1、DNMT3a 和 DNMT3b 催化的。转录控制区的 DNA 甲基化常常与转录沉默状态有关，因为 DNA 甲基化涉及沉默的建立和维持。由于这个原因，在基因座或目的控制区内的胞嘧啶甲基化状态的检查常常是转录调控分析中的一个重要步骤。

为特定基因组位点胞嘧啶甲基化的监测所发展的第一种方法，是用甲基化敏感性限制性内切核酸酶切割基因组 DNA，之后使用检测目的基因组区的探针进行 Southern 印迹分析（Shemer et al. 1991；Kafri et al. 1992）。也可以使用两条位于切割位点两侧的引物，通过 PCR 和实时-PCR 分析甲基化敏感性限制性内切核酸酶产生的切割产物（如见 Mostoslavsky et al. 1998）。

尽管在早期研究中上述策略提供了宝贵的认识，但主要的局限性是该策略只能用于研究落入甲基化敏感性限制性内切核酸酶识别序列中的 CpG 二核苷酸的状态。为了获得基因或控制区内部关于胞嘧啶甲基化状态的全面信息，又发展了亚硫酸氢盐测序技术（Frommer et al. 1992；Clark et al. 1994；Clark et al. 2006），该技术始于纯化的基因组 DNA 的亚硫酸氢盐处理。在合适的反应条件下，亚硫酸氢盐将胞嘧啶转变为尿嘧啶，但是 5-甲基胞嘧啶保持无反应。亚硫酸氢盐处理之后，通过 PCR 扩增目的 DNA 区域，并对 PCR 产物进行测序（一般在测序前，先将 PCR 产物克隆进质粒，再测序）。通过与已知 DNA 序列相比较，可鉴定基因组 DNA 区域内的甲基化胞嘧啶，因为它们将是仅有的未被转变成尿嘧啶的胞嘧啶。换句话说，基因组 DNA 中甲基化的 CG 碱基对在 DNA 序列中将仍然是 CG 碱基对，而非甲基化的 CG 碱基对将变成 TA 碱基对。可以从文献中找到非常好的亚硫酸氢盐测序方案（如 Clark et al，1994，2006 ），他们还详细描述了在结果解释中可能遇到的困难和重要对照。

对于获得目的区域胞嘧啶甲基化的详细信息而言，亚硫酸氢盐测序技术仍然是首选策略。然而，已经发展了很多不太费力并且可被修改用于全基因组研究的替代技术（见综述 Suzuki and Bird 2008），限制性界标基因组扫描（RLGS）就是其中之一（Costello et al. 2002；Yu et al. 2005），该技术已经非常成功地用于许多研究中，但是该技术受到对甲基化敏感性内切核酸酶依赖的限制。另一项技术，即甲基化 DNA 免疫沉淀（mDIP）（ Weber et al. 2005 ）已经被更加广泛地用于各种细胞类型和物种，包括拟南芥和哺乳动物肿瘤细胞系及胚胎干细胞的 DNA 甲基化的全基因组研究（Keshit et al. 2006；Zhang et al. 2006；Fouse et al. 2008）。mDIP 技术利用免疫沉淀反应实验中的 5-甲基胞苷抗体从非甲基化 DNA 片段中分离甲基化 DNA 片段，分离的甲基化 DNA 片段之后可与全基因组叠连阵列或启动子阵列进行杂交。这种分析的局限性是它的分辨率

中等，这是因为5-甲基胞苷抗体不能区分所有胞嘧啶甲基化的片段和仅有部分胞嘧啶甲基化的片段。在哺乳动物细胞中，这种方法对于研究CpG岛的DNA甲基化最有用。

为了获得全基因组水平上DNA甲基化详尽而全面的图景，全基因组亚硫酸氢盐测序是必需的。通过亚硫酸氢盐测序和大规模平行测序的结合，最近在拟南芥基因组中实现了这个目标（Cokus et al. 2008），并且在小鼠胚胎干细胞基因组的部分分析也实现了这个目标（Meissner et al. 2008）。

技术

方案9.1　MNase-Southern印迹分析

概要

微球菌核酸酶（MNase）在核酸酶中的独特之处在于它能诱导核小体连接区内双链的断裂，而在核小体本身内部仅造成单链切刻（nick）的相对能力。由于这种特性，MNase可用于确定目的DNA片段是否是核小体DNA。另外，如果核小体定位一致，MNase可用于确定DNA区域上核小体的大致位置（见专题9.2）。简单地说，首先裂解细胞，通过离心分离细胞核，然后在细胞核中加入极限浓度的MNase，从而引起核小体连接区上的切割（最好在每个DNA分子的两个位点上切割）。终止切割反应，纯化基因组DNA。纯化的DNA经琼脂糖凝胶电泳、溴化乙锭染色后，应该产生大小对应于核小体核心长度加连接间区长度（约200 bp）的倍数的梯状条带。为确定目的DNA片段是否是核小体DNA，可以通过Southern印迹对基因组DNA进行分析。如果来源于DNA片段的探针与核小体梯状条带相杂交，该片段可能确实被包装进了核小体。为了确定核小体定位，必须在凝胶电泳和Southern印迹分析之前用限制酶切割纯化的基因组DNA（见专题9.2）。下文介绍的方案改编自Richard-Foy和Hager（1987）及Enver等（1985）的方案。

时线与安排

在进行该方案之前，除下文列出的试剂以外，需要的特殊试剂仅是细胞和用于制备Southern印迹探针的DNA片段。有了这些试剂，完成这个完整的方案将需数天。第一天，配制缓冲液，收集细胞，制备细胞核，然后用MNase消化细胞核，接着用蛋白酶K孵育过夜，降解MNase和细胞核蛋白。第二天纯化基因组DNA，用合适的限制酶切割数小时或过夜，然后进行Southern印迹，包括凝胶电泳、转膜及放射性标记探针杂交。洗膜后，在胶片或磷屏成像屏上进行曝光。

提纲

低分辨率体内MNase分析：

步骤1：配制缓冲液（1 h）。

步骤2：收集细胞、制备细胞核（1 h）。

步骤 3：以 MNase 消化（孵育过夜后，需 5 min）

步骤 4：纯化基因组 DNA（1 天）。

步骤 5：以限制酶体外消化纯化的 MNase 处理 DNA（过夜）。

步骤 6：琼脂糖凝胶电泳和 Southern 印迹转移（1 天）。

步骤 7：制备放射性标记的 Southern 探针（1 h）。

步骤 8：Southern 印迹预杂交和杂交（1 天）。

步骤 9：洗膜，在胶片或磷屏成像仪上曝光（2～4 h 至过夜）。

步骤 10：（可选）剥去膜上放射性探针（1 h）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

琼脂糖凝胶（1 %）

$CaCl_2$（1 mmol/L）〈!〉

细胞（10^8 个）

氯仿〈!〉

去离子水

变性液

  0.5 mol/L NaOH〈!〉

  1.5 mol/L NaCl

  使用前新鲜配制

20×Denhardt's 溶液

  2 g 聚蔗糖

  2 g 聚乙烯吡咯烷酮（PVP）

  2 g BSA（牛血清白蛋白）（Pentax 组分 V）

  加 dH_2O 至 500 ml

  通过 0.45 μm 滤膜过滤。－20℃储存。

蒸馏水

用于准备 Southern 印迹的 DNA 片段

DNA 分子质量标准物

乙醇〈!〉

溴化乙锭〈!〉

甲酰胺（可选）〈!〉

6×凝胶上样缓冲液

基因组 DNA

杂交液（每张膜）

  0.4 ml　3 mg/ml 超声波处理的变性鲑鱼精 DNA

4.6 ml H_2O
1.2 ml 10 % SDS〈!〉
3 ml 40%硫酸葡聚糖钠
0.6 ml 20×Denhardt's 溶液
2.4 ml 20×SSC
使用前新鲜配制。

微球菌核酸酶（MNase）

MNase 消化缓冲液
10 mmol/L Tris-HCl（pH 7.4）
15 mmol/L NaCl
60 mmol/L KCl〈!〉
0.15 mmol/L 精胺〈!〉
0.5 mmol/L 亚精胺〈!〉
4℃保存。

MNase 终止缓冲液
100 mmol/L EDTA
10 mmol/L EGTA
调至 pH 7.5。室温保存。

NP-40 裂解缓冲液
10 mmol/L Tris-HCl（pH 7.4）
10 mmol/L NaCl
3 mmol/L $MgCl_2$〈!〉
0.5 % NP-40（Nonidet P-40）〈!〉
0.15 mmol/L 精胺〈!〉
0.5 mmol/L 亚精胺〈!〉
4℃保存

中和液
1 mol/L Tris-HCl（pH 8）
1.5 mol/L NaCl
使用前新鲜配制。

苯酚：氯仿（pH 8）〈!〉

1×磷酸缓冲盐

预杂交缓冲液
4×SSC
1×Denhardt's 溶液
1% SDS〈!〉
100 μg/ml 超声波处理的变性鲑鱼精 DNA
临用前新鲜配制

蛋白酶 K（25 mg/ml）〈!〉
放射性标记探针〈!〉
限制性内切核酸酶
RNase A（热处理的）〈!〉
SDS（20%）〈!〉
乙酸钠（3 mol/L，pH 5.2）
6×SSC
20×SSC
175.3 g NaCl
88.2 g 柠檬酸钠〈!〉
用 dH_2O 调至 1L。
调节 pH 到 7.2。室温保存。
剥离液（可选）
0.2 mol/L NaOH〈!〉
0.1×SSC
1% SDS〈!〉
TE
10 mmol/L Tris-HCl（pH 7.5）
1 mmol/L EDTA
转印缓冲液（10×或 20×SSC）
洗涤缓冲液 1：2×SSC，0.1% SDS〈!〉
洗涤缓冲液 2：1×SSC，0.1% SDS〈!〉
洗涤缓冲液 3：0.5×SSC，0.1% SDS〈!〉
洗涤缓冲液 4：0.2×SSC，0.1% SDS〈!〉

设备

毛细转印装置
离心机
锥形管（50 ml）
带盖容器
干燥器（可选）
盖革计数器
凝胶电泳仪
杂交袋（可封口）
杂交炉
冰
培养箱
微量离心机

微量离心管
用于 Southern 印迹的尼龙膜
巴斯德吸管
磷屏成像仪或胶片
保鲜膜
推式柱（Stratagene 400701 NucTrap 探针纯化柱）
剃刀刀片
直尺
振荡器
分光光度计
真空离心蒸发浓缩器（SpeedVac）
UV 灯箱〈!〉
水浴
Whatman 滤纸

程序

步骤 1：配制缓冲液

步骤 2：收集细胞并制备细胞核

以大约 10^8 个细胞为起始材料，用于多达 9 个不同的 MNase 消化条件。

1. 在一次性 50 ml 锥形管中，4℃，1500 r/min 离心 10 min，沉淀细胞。

2. 弃上清，用冰冷的 1×PBS（10 ml）洗涤细胞沉淀，4℃，1500 r/min 离心 10 min，沉淀细胞。

3. 用 5 ml 冰冷的 NP-40 裂解缓冲液重悬细胞沉淀，冰上孵育 5 min。

整个程序中，细胞核都应保持冰冷。进行操作及向离心机转移时都应置于冰上。

4. 4℃，1000 r/min 离心 10 min 沉淀细胞核，弃上清。

细胞核会形成疏松的沉淀。在除去上清时注意不要弄散沉淀，倒掉上清或用巴斯德吸管吸除。从此步往后不要真空吸液。

步骤 3：以 MNase 消化

1. 用 2.5 ml MNase 消化缓冲液洗涤细胞核。

2. 4℃，1000 r/min 离心 10min，沉淀细胞核。小心弃上清，用 1 ml 含有 1 mmol/L $CaCl_2$ 的 MNase 消化缓冲液重悬细胞核。

每 1 ml MNase 消化缓冲液中加入 1μl 1mol/L $CaCl_2$。

3. 在一系列含有稀释的 MNase 的微量离心管中各加入 100μl 细胞核。

应该对 MNase 进行滴定。作为起始范围，使用 0.25～72 U。必须根据滴定结果优化 MNase 浓度。

4. 样品于室温下孵育 5 min。

5. 每个样品中加入 80 μl MNase 消化缓冲液和 20 μl MNase 终止缓冲液。

6. 加入 3 μl 蛋白酶 K（25mg/ml）和 10 μl 20% SDS。

步骤 5 和步骤 6 终止 MNase 消化反应，应迅速进行。

7. 37 ℃ 孵育过夜。

步骤 4：纯化基因组 DNA

1. 用 200 μl 酚∶氯仿（pH 8.0）抽提样品。

为了防止基因组 DNA 样品的广泛剪切，不要涡旋振荡样品，而改为通过轻弹管壁或振荡来混匀。

2. 在微量离心机中高速离心样品 5 min，将水相小心转移到新微量离心管。基因组 DNA 可能会表现为不溶性的“线”状。用移液管缓慢吸移水相和水相中的“线”状 DNA，尽量减小对酚∶氯仿界面的扰动。

3. 用 200 μl 氯仿抽提样品。采取与步骤 1 相同的预防措施。

4. 在微量离心机中高速离心样品 5 min，将水相小心转移至新的微量离心管。采取与步骤 2 相同的预防措施。

5. 每个样品加入 2 μl 热处理的 RNase A（10 mg/ml），37℃孵育 2 h。

该步骤去除基因组 DNA 制剂中的 RNA。

6. 用 200 μl 酚∶氯仿（pH 8.0）抽提样品。采取与步骤 1 相同的预防措施。

7. 在微量离心机中高速离心样品 5 min，将水相小心转移至新的微量离心管。采取与步骤 2 相同的预防措施。

在此阶段，界面通常不清晰，要避免吸取界面。如果吸进少量时，可以在氯仿抽提时去除。

8. 用 200 μl 氯仿抽提样品。采取与步骤 1 相同的预防措施。

9. 在微量离心机中高速离心样品 5 min，将水相小心转移至新的微量离心管。这时不要吸取任何不清晰的界面，最好稍留一些水相。

10. 加入 1/10 体积的 3 mol/L 乙酸钠（pH 5.2）和 2.5 体积的冰冷乙醇，沉淀消化的基因组 DNA，将样品置于－20℃孵育 1 h。

11. 在微量离心机中 4℃高速离心 10 min，沉淀 DNA。

12. 弃上清，用冰冷的 70%乙醇洗沉淀，4℃高速离心样品 1～2 min。

13. 弃上清，在操作台面上或利用无热 SpeedVac 干燥 DNA。

注意不要使 DNA 过于干燥，否则难以溶解。

14. 将沉淀轻轻重悬于 100 μl H_2O 或 TE（10 mmol/L Tris-HCl，pH 7.5，1 mmol/L EDTA）中。

不要涡旋或剧烈吹打，可能需要使 DNA 在 4℃环境下溶解过夜。

步骤 5：以限制性内切核酸酶体外消化纯化的 MNase 处理 DNA

测试目的区域周围的多个限制性内切核酸酶位点。推荐一个核小体（147 bp）距离

内的邻近限制性内切核酸酶切割位点。

1. 使用分光光度计确定 DNA 浓度。5～20 μg DNA 用于体外限制性内切核酸酶消化分析。

如果 RNaseA 消化不彻底，则计算的 DNA 浓度可能有误导性。

2. 37℃过夜消化 35 μl 总体积中的 DNA。在其最适缓冲液中使用过量的限制性内切核酸酶，以确保彻底消化。

步骤 6：琼脂糖凝胶电泳和 Southern 印迹转移

1. 灌制长约 25 cm、含溴化乙锭的 1%的琼脂糖凝胶（见 Sambrook and Russell 2001，5.4）。

百分浓度较高的琼脂糖（如 1.4%）能提高分辨率，但是也会降低 DNA 转移的效率。

2. 在过夜消化的限制性内切核酸酶酶切产物中加入 7 μl 6×凝胶上样缓冲液（II 型，Sambrook and Russell 2001，5.42）至终浓度为 1×。

3. 将全部样品（42 μl）加到琼脂糖凝胶的一个上样孔中。在相邻的上样孔中加入 DNA 分子质量标准物。

4. 以低于 1 V/cm 的电压电泳。

为了得到 DNA 片段更紧凑的迁移效果，琼脂糖凝胶要缓慢电泳。

5. 在旁边带有标尺的 UV 灯箱上进行琼脂糖凝胶成像。用剃刀切去点样孔，确保保留一个参照用于之后 DNA 片段大小的估计。

分子质量标准物会强烈地非特异性杂交放射性标记探针，从而掩盖相邻的样品泳道。因此，用剃刀去除分子质量标准物泳道是有用的。

6. 用短波 UV 照射凝胶 2～5 min。

7. 将凝胶浸泡于变性液中在摇床上摇 45～60 min。

每次都新鲜配制变性液。

8. 去除变性液，用去离子水短暂淋洗。将凝胶浸泡于中和液中在摇床上摇 45～60 min。

9. 准备用于转印的尼龙膜。先将膜剪成凝胶大小，为了便于确定方向，在膜和凝胶的相应位置上切一槽口。将膜在去离子水中浸泡 5 min，然后将膜放在转印缓冲液（10×或 20×SSC）中浸泡至少 5 min。

推荐使用尼龙膜，特别是带有正电荷的尼龙膜。不推荐硝酸纤维素膜，因为硝酸纤维素膜具有一些缺点（Ausubel et al. 1994，pp. 2.9.11-2.9.12；Sambrook and Russell 2001，6.37）。

10. 按 Ausubel 等（1994，pp. 2.9.7-2.9.8）和 Sambrook 及 Russell（2001，6.34）所述方法准备毛细转移装置。然后按前面（Ausubel et al. 1994，pp. 2.9.7-2.9.8；Sambrook and Russell 2001，6.39）所述的方法在中性条件下进行过夜毛细转印，将 DNA 转印至尼龙膜上。

可通过溴化乙锭再染胶和确定胶中仍然保留的 DNA 量，间接确定转印的效率。

11. 按 Ausubel 等（1994，pp. 2. 9. 5～2. 9. 6）或 Sambrook 和 Russell（2001，6. 46）所述，将 DNA 通过 UV 通过交联（254nm）到尼龙膜上。

完成 UV 交联步骤后，可以保存膜。短期保存时，将干燥的膜夹在 Whatman 滤纸之间，保存于室温。长期保存时，将膜夹在 Whatman 滤纸之间，在干燥器中 4℃保存。

步骤 7：制备放射性标记的 Southern 探针

当设计 DNA 探针时，应考虑以下几点。

- 理想的情况下，探针应设计在用于体外消化的限制性内切核酸酶位点的一端。如果同一张印迹膜上分析多个限制性内切核酸酶消化产物，探针应该只杂交到每个限制性内切核酸酶位点的一侧。
- 为了确保探针只会检测与体外限制性内切核酸酶切割相关的 MNase 消化片段，探针的大小应限制在近似一个核小体和连接区的长度（约 200 bp）。

1. 可以通过切刻平移法（见 Ausubel et al. 1994，pp. 3. 5. 4-3. 5. 6；Sambrook and Russell 2001，9. 12）或随机引物法（Ausubel et al. 1994，pp. 3. 5. 9-3. 5. 10；Sambrook and Russell 2001，9. 4）制备放射性标记探针。

2. 从游离核苷酸中纯化放射性标记探针。

可以通过使用 Stratagene 推式柱（400701，NucTrap 柱；见方案 3. 1）或相似方法实现这一点。

步骤 8：Southern 印迹预杂交和杂交

1. 将膜在 6×SSC 中浸泡 2 min，浸入前首先让膜漂浮在 6×SSC 表面。

2. 将膜置于可封口的杂交袋（如果杂交在水浴里进行）或杂交管（如果使用杂交炉）中。用预杂交液完全将膜覆盖。如果使用可封口的杂交袋，则封口时要排除气泡（气泡会影响预杂交和杂交）。

可以使用以甲酰胺制备的替代预杂交液（见 Sambrook and Russell 2001，6. 50）。

3. 65℃预杂交 2 h。如果使用甲酰胺预杂交液，则 42℃预杂交 2 h。

4. 配制杂交液（参见步骤 1）。

关于替代杂交缓冲液的成分（含或不含甲酰胺），见 Ausubel 等（1994，p. 2. 10. 7）及 Sambrook 和 Russell（2001，6. 50）。

5. 如果使用双链 DNA 探针，将探针置于沸水浴中 5 min。

每 10 ml 杂交液使用（40～100）$\times 10^6$ cpm 放射性标记探针。

6. 将探针简单离心收集管壁上冷凝液滴，并迅速置于冰上。

7. 除去预杂交液，加入 10 ml 杂交液。在杂交液中加入煮过的、放射性标记的探针，小心封口。将内容物混匀，以保证探针均匀分布。

如果使用可封口的袋子，封口时小心不要留有气泡。

8. 65℃杂交过夜。

杂交温度根据特定探针的不同有所变动。如果使用含有甲酰胺的杂交缓冲液，杂交在 42℃进行。

步骤 9：洗膜、在胶片或磷屏成像仪上曝光

1. 小心弃除杂交液，用洗涤缓冲液 1（2×SSC，0.1% SDS）室温下洗膜 5 min。

在带有盖子的容器中进行漂洗很适宜。将容器置于摇床上轻摇。对具有放射性的液体/固体进行操作时要适当小心。

2. 弃除洗涤缓冲液 1，加入新鲜的洗涤缓冲液 1。室温下在摇床上孵育 15 min。

监测洗涤缓冲液的放射性。第一次洗涤缓冲液可能含有大量的放射性。

3. 弃除洗涤缓冲液 1，加入洗涤缓冲液 2（1×SSC，0.1% SDS）。室温下在摇床上孵育 15 min。

随着 SSC 浓度的减小，洗涤的严紧性增加。这将去除非特异性杂交的探针。

4. 弃除洗涤缓冲液 2，加入洗涤缓冲液 3（0.5×SSC，0.1% SDS）。室温下在摇床上孵育 15min。

5. 弃除洗涤缓冲液 3，监测膜的放射性。

随着洗涤严紧性的增加，这有助于监测膜，以确定放射计数丢失的速率。

6. 如果检测到膜上有显著的放射量，加入洗涤缓冲液 4（0.2×SSC，0.1% SDS），室温下在摇床上孵育 15 min。

7. 弃除洗涤缓冲液，监测膜的放射性。如果保留显著的放射量，加入洗涤缓冲液 2，42℃下孵育 15 min。

随着温度的增加，洗涤的严紧性也增加。在温度增加的洗涤过程中，要小心监测膜。

8. 如果需要，以 42℃孵育重复步骤 4～6。

9. 如果仍有显著的放射性，以 65℃孵育重复步骤 3～6。

10. 通过 Whatman 滤纸吸收除去杂交膜上多余的液体。保鲜膜包住膜，在胶片或磷屏成像仪上曝光。

整个过程不允许膜完全变干。

步骤 10（可选）：剥膜

1. 将印迹膜置于剥离液（0.2 mol/L NaOH，0.1×SSC，1% SDS）中。65℃孵育 30 min。

这一步将从膜上去除杂交的放射性探针，然后可以用不同的探针进行杂交。

2. 用蒸馏水洗膜。

3. 用含有 4×SSC 和 1% SDS 的溶液洗膜两次。

4. 膜可以进行再一次的杂交，或保存备用（见上文）。

附加说明

1. 将限制性内切核酸酶未体外消化的 MNase 消化 DNA 样品和限制性内切核酸酶消化的样品平行电泳是有益的。如果目的 DNA 序列包含在核小体当中，该样品将会表现为对应于核小体核心加连接区的倍数的梯带。该样品也会显示出标准核小体梯带的位置，以便与用限制性内切核酸酶消化 DNA 得到的梯带相比较。体外限制性内切核酸酶

消化时梯带的保留暗示核小体特异性地定位在该 DNA 区域中。与未酶切的梯带相比，体外限制性内切核酸酶消化时梯带的迁移暗示该限制性内切核酸酶位点位于定位的核小体中（见专题 9.2）。

2. 使用跨越超过一个核小体的探针通常更为适宜。如果情况是这样或者探针与限制性内切核酸酶酶切位点不直接毗邻，就必须更谨慎地解释数据。比如，杂交结果会产生两种截然不同的核小体梯状条带：一个与未切割的 MNase 消化 DNA 梯带一致，一个和它相比有所迁移。发生这种现象的原因是限制性内切核酸酶切割位点仅位于和探针杂交的一部分 DNA 片段的一端。

3. 可以使用与这与个同样的程序进行 DNase I 敏感性和 DNase I 高敏感性研究，只是将方案 9.2 中所述的 DNase I 切割步骤替换为这里所述的 MNase 切割步骤。

疑难解答

弱杂交或无杂交

可能原因：杂交或洗膜条件过于严紧。

解决办法：通常有必要凭经验确定杂交条件。较低杂交温度会减小严紧性。也可能在洗膜步骤中特异性杂交丢失。在洗膜过程中监测膜上的放射性，当保留在膜上的放射性减弱时，停止洗膜步骤。

可能原因：DNA 转移至膜上效率低。

解决办法：监测转印的一个方法是确定转膜后琼脂糖凝胶上是否还存在有 DNA。如果还保留有高比例的 DNA，则转膜效率较低。即使琼脂糖凝胶上不再能检测到 DNA，也不一定表明转膜效率高。可以使用亚甲蓝将膜染色（Ausubel et al. 1994，p. 2. 10. 14），也可以考虑使用不同的印迹技术缓解转膜问题（见 Ausubel et al. 1994，单元 2. 9；Sambrook and Russell 2001，6. 33）。

可能原因：探针标记效率低。

解决办法：基因组 Southern 印迹通常需要高比活性的探针。探针的质量对于获得结果很关键。探针应该“通体被标记”以增强比活性。不推荐末端标记的探针。

高背景

可能原因：杂交条件不够严紧。

解决办法：探针可能非特异性地与 DNA 杂交。在这种情况下，提高杂交温度可能有助于解决该问题。或者，通过使用 50%甲酰胺缓冲液进行杂交，以提高杂交的严紧性（关于缓冲液成分，见 Ausubel et al. 1994，p. 2. 10. 7；Sambrook and Russell 2001，6. 50）。

可能原因：洗涤不够严紧。

解决办法：通过提高温度、降低 SSC 浓度、增加 SDS 浓度，提高洗涤的严紧性。在更加严紧的条件下进行洗涤。

可能原因：预杂交/杂交封闭效率低。

解决办法：所使用的封闭试剂（Denhardt's 溶液、鲑鱼精 DNA 或其他）很关键，必须是高质量的。因此，可以测试新的试剂，还应该延长预杂交的时间，或考虑使用其他的封闭试剂（见 Ausubel et al. 1994，p. 2. 10. 16；Sambrook and Russell 2001，6. 50）。

其他疑难解答

在 Southern 印迹技术中，许多地方都会出现问题。关于可能出现的问题和解决办法的更全面的罗列，见 Ausubel 等（1994，单元 2. 10）及 Sambrook 和 Russell（2001，6. 56）。

方案 9. 2　LM-PCR 方法

概要

本方案提供下列分析的操作程序：DNase I 基因组足迹、核小体定位的 MNase 作图、监测核小体重塑的限制性内切核酸酶可及性分析及 DMS 基因组足迹。

LM-PCR 方法是 Wold 及其同事在 1988 年（Mueller et al. 1988）发展的，它可以作为一种检测复杂样品中 DNA 链断裂的灵敏技术（见专题 9. 1）。已经证明这一通用方法对于许多研究目的，特别是分析内源 DNA 基因座特征具有重要价值。如正文及下述方案所述，LM-PCR 可以与许多 DNA 修饰和切割试剂联用。LM-PCR 与 DNase I 消化或 DMS 修饰联用产生类似体外 DNase I 足迹或甲基化保护的分析方法（见第 11 章）。LM-PCR 与 MNase 消化联用可进行核小体定位的高分辨率分析。LM-PCR 与高锰酸钾修饰联用可揭示熔解的 DNA 区域。LM-PCR 与分离的细胞核中的限制性内切核酸酶消化联用可为核小体重塑提供证据。这些分析方法，除高锰酸盐修饰法外，都将在下文进行介绍。

简单地说，为了进行 DNase I、MNase 和限制性内切核酸酶分析，首先必须通过利用非离子去污剂裂解细胞，离心分离细胞核。然后，用合适的核酸酶处理分离的细胞核，接着纯化基因组 DNA。和上面不同的是，DMS 直接加到完整的细胞上。终止修饰反应后，纯化基因组 DNA，通过哌啶诱导处理，在被修饰的碱基附近切割。然后，对通过上述任何一种方法切割得到的基因组 DNA 进行 LM-PCR。关于 LM-PCR 策略的一般概要，见图 9. 4。

下文描述的方案来源于多种来源。Wold 及其同事发展了 LM-PCR 方法（Mueller et al. 1988；Garrity and Wold 1992；Ausubel et al. 1994，Chapter15. 5；Garrity et al. 1995），此处的 LM-PCR 方法是该方法的一个修改版。DNase I 方法源自于 Enver 等（1985）。MNase-LM-PCR 方法首先由 McPherson 等（1993）发展，这里描述的方案还包括来自 Richard-Foy 和 Hager（1987）的 MNase 消化步骤。限制性内切核酸酶方法修改自文献 Reik 等（1991），DMS 方法来自 Mueller 等（1988）和 Ausubel 等（1994）。关于 DMS 方案更详细的描述，强烈建议参考 Ausubel 等（1994，单元 15. 5）。

时线与安排

连接介导 PCR（LM-PCR）通常需要使用三个基因特异性的引物和两个能退火形成双链接头的互补寡核苷酸，其中一个接头寡核苷酸也用作 PCR 步骤中的接头引物。尽管体内消化所产生的序列具有可变性，但接头提供恒定的 5′ 序列。接头的制备需要过夜孵育，因此，在进行这个方案的 LM-PCR 部分之前应制备接头。步骤 6（LM-PCR，第 1 天）中含有接头制备的操作指南，但应在此步之前完成制备。

基因特异性引物的设计和完整性是 LM-PCR 程序成功的关键。考虑以下几点指导方针将对实验有所帮助。

- 基因特异性引物 1 和引物 2 最好为 25 mer，引物 3 为 30 mer。
- 基因特异性引物的 G/C 含量应随引物系列号的增加而增加。G/C 含量增加会导致 T_m 的增加。G/C 含量为 48%（引物 1）、56%（引物 2）、53%～60%（引物 3）时效果好。
- 引物 3 应该与距离目的区 50～200bp 的部位杂交（对于限制性内切核酸酶可及性分析而言，应为 50～350bp）。
- 接头和基因特异性引物都应凝胶纯化，确保引物的全长和纯度。
- 理想情况是引物 3 和引物 2 至少有一半序列相重叠。据认为在标记循环过程中，用重叠序列能更有效地排除引物 2。
- 当计算最终的 LM-PCR 产物大小时，记住须加上 25 bp，以解释连接的接头序列。

提纲

体内 DNase I、MNase 和限制性内切核酸酶足迹：

步骤 1：配制缓冲液（2 h）。

步骤 2：收集细胞并制备细胞核（1 h）。

步骤 3：使用核酸酶消化（15 min 及过夜）。

步骤 4：纯化 DNA（1 天）。

步骤 5：制备用于 LM-PCR 的 DNA（2～4 h）。

步骤 6：LM-PCR 第 1 天（1 天）。

步骤 7：LM-PCR 第 2 天（1 天）。

体内 DMS 足迹：

步骤 1：配制缓冲液（2 h）。

步骤 2：用 DMS 处理细胞并收集细胞（5 h）。

步骤 3：纯化 DNA（1 天）。

步骤 4：用哌啶处理（4 h）。

步骤 5：制备用于 LM-PCR 的 DNA（2～4 h）。

步骤 6：LM-PCR 第#1 天（1 天）。

步骤 7：LM-PCR 第#2 天（1 天）。

体内 DNase I、MNase 及限制性内切核酸酶足迹

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

琼脂糖凝胶
5×扩增缓冲液
　　200 mmol/L NaCl
　　100 mmol/L Tris-HCl（pH 8.9）
　　25 mmol/L $MgSO_4$〈!〉
　　0.05% 明胶
　　0.5% Triton X-100〈!〉
　　−20℃保存。
ATP（1 mmol/L 和 100 mmol/L）
[γ-^{32}P] ATP（6000 Ci/ mmol）〈!〉
牛血清白蛋白（无 DNase；10 mg/ml）
溴酚蓝染料〈!〉
缓冲液 A（DNase I 消化缓冲液）
　　100 mmol/L NaCl
　　50 mmol/L Tris-HCl（pH 8.0）
　　3 mmol/L $MgCl_2$〈!〉
　　0.15 mmol/L 精胺〈!〉
　　0.5 mmol/L 亚精胺〈!〉
　　4℃保存。
$CaCl_2$（1 mol/L 和 1 mmol/L）〈!〉
细胞裂解和酶消化试剂
细胞（每个反应 3×10^6～4×10^6 个）
氯仿〈!〉
二硫苏糖醇（DTT，0.5 mol/L）〈!〉
DNA（纯化的基因组 DNA 或质粒基因组克隆 DNA；用于对照的 DNA）
DNase I（2.5 mg/ml）
dNTP 混合物（25 mmol/L）
　　25 mmol/L dATP
　　25 mmol/L dCTP
　　25 mmol/L dGTP
　　25 mmol/L dTTP

-20℃保存。

EDTA（0.5 mol/L）

乙醇（冰冷）〈!〉

5×第一链缓冲液

200 mmol/L NaCl

50 mmol/L Tris-HCl（pH 8.9）

25 mmol/L $MgSO_4$〈!〉

0.05 %明胶

-20℃保存。

甲酰胺上样缓冲液

80 %（*m/V*）去离子甲酰胺〈!〉

10 mmol/L EDTA（pH 8.0）

1 mg/ml 二甲苯青〈!〉FF

1 mg/ml 溴酚蓝〈!〉

LM-PCR 试剂

接头寡核苷酸（LM-PCR 1）

（5′→3′）：GCGGTGACCCGGGAGATCTGAATTC

接头寡核苷酸（LM-PCR 2）

（5′→3′）：GAATTCAGATC

$MgCl_2$（1 mol/L）〈!〉

微球菌核酸酶（MNase）

微球菌核酸酶（MNase）消化缓冲液

10 mmol/L Tris-HCl（pH 7.4）

15 mmol/L NaCl

60 mmol/L KCl〈!〉

0.15 mmol/L 精胺〈!〉

0.5 mmol/L 亚精胺〈!〉

4℃保存。

MNase 终止缓冲液

100 mmol/L EDTA

10 mmol/L EGTA

调节 pH 至 7.5。室温保存。

NP-40 裂解缓冲液

10 mmol/L Tris-HCl（pH 7.4）

10 mmol/L NaCl

3 mmol/L $MgCl_2$〈!〉

0.5 % NP-40（Nonidet P-40）〈!〉

0.15 mmol/L 精胺〈!〉

0.5 mmol/L 亚精胺〈!〉
4℃保存。
pfu 聚合酶（0.25 μl）
苯酚：氯仿（pH 8）〈!〉
1×磷酸缓冲盐
聚丙烯酰胺凝胶（变性，8%）
蛋白酶 K（25 mg/ml）〈!〉
2×蛋白酶 K 缓冲液
100 mmol/L Tris-HCl（pH 7.5）
200 mmol/L NaCl
2 mmol/L EDTA
1% SDS〈!〉
室温保存。
限制性内切核酸酶消化缓冲液
10 mmol/L Tris-HCl（pH 7.4）
50 mmol/L NaCl
10 mmol/L $MgCl_2$〈!〉
0.2 mmol/L EDTA
0.2 mmol/L EGTA
0.15 mmol/L 精胺〈!〉
0.5 mmol/L 亚精胺〈!〉
1 mmol/L β-巯基乙醇〈!〉
4℃保存。
限制性内切核酸酶
终止液
10 mmol/L Tris-HCl（pH 7.5）
4 mmol/L EDTA
260 mmol/L 乙酸钠（pH 7.0）
67 μg/ml 酵母 tRNA
使用前新鲜配制。
RNaseA（热处理；10 mg/ml）〈!〉
SDS（20%）〈!〉
乙酸钠（3 mol/L，pH 5.2 和 3 mol/L，pH 7.0）
T4 DNA 连接酶（1.0 μl，3 U）
T4 多核苷酸激酶（10 U/ml；New England Biolabs）
10×T4 多核苷酸激酶缓冲液
700 mmol/L Tris-Cl（pH 7.6）
100 mmol/L $MgCl_2$〈!〉

50 mmol/L DTT 〈!〉
小量分装，−20℃保存。

TE

10 mmol/L Tris-HCl（pH 7.5）
1 mmol/L EDTA

Tris-HCl（pH 7.7 和 1 mol/L，pH 7.5）

酵母 tRNA（10 mg/ml）

设备

离心机
锥形管
DNA 测序胶
凝胶电泳仪
加热块
高效液相色谱（HPLC）仪
冰
培养箱
盖锁（LidLok）
LM-PCR 仪
微量离心机
微量离心管
PCR 管
磷屏成像仪
移液器
推式柱（Stratagene 400701 NucTrap 探针纯化柱）
放射性防护罩
分光光度计
真空离心蒸发浓缩器（SpeedVac）

程序

步骤 1：配制缓冲液

步骤 2：收集细胞并制备细胞核

从每个消化反应 3×10^6～4×10^6个细胞开始。该细胞核消化程序可根据所需样品数量进行缩放，这里显示的程序需要 15×10^6～20×10^6 细胞（5 个消化反应）。

1. 在一次性锥形管中，4℃，1500 r/min 离心 10 min，沉淀细胞。
2. 弃上清。用冰冷的 1×PBS 洗涤细胞沉淀。4℃，1500 r/min 离心 10 min，沉淀细胞。

3. 用 2.5 ml 的 NP-40 裂解缓冲液（冰冷）重悬细胞沉淀，冰上孵育 5 min。

缓冲液中所包含的多价阳离子精胺、亚精胺用于防止细胞核结块。重要的是在其余的所有程序中要保持样品冰冷；在操作或向离心机转移等过程中应置于冰上，所用的随后的缓冲液都应是冰冷的。

4. 4℃，1000 r/min 离心 10min，沉淀细胞核。弃上清。

细胞核将形成疏松的沉淀。小心除去上清，从此步往后不要使用真空移液器。

步骤 3：核酸酶消化

制备好细胞核之后，该程序根据特定的消化类型而异。这里概述了每种类型的消化：（A）DNase I；（B）MNase；（C）限制性内切核酸酶。本方案末尾还包含了 DMS 程序，该程序涉及完整细胞的 DMS 处理。

DNase I 消化

1A. 用 1.3 ml 的缓冲液 A 洗涤细胞核沉淀。

2A. 4℃，1000 r/min 离心 10 min，沉淀细胞核。小心弃上清，用 500 μl 含 1 mmol/L $CaCl_2$ 的缓冲液 A 重悬细胞核。

每 1 ml 缓冲液 A 中加入 1μl 1 mol/L $CaCl_2$。钙离子为 DNase I 活性所需。

3A. 转移 100 μl 细胞核至含有稀释 DNase I 的各离心管中。

DNase I 应分成小份，以 2.5 mg/ml 的浓度于－80℃保存。使用前，将 DNase I 稀释至终浓度 0.15 μg/μl。作为起始点，可以尝试从 0.15 μg/反应到 1.5 μg/反应的滴定范围。必须根据 LM-PCR 结果优化 DNase I 用量。

4A. 将样品于 37℃孵育 2 min。

DNase I 消化时间与温度可能需要优化。

5A. 每个样品中加入 2 μl 0.5mol/L EDTA 和 100 μl 缓冲液 A。

EDTA 螯合钙离子和镁离子。

6A. 每个样品中加入 3 μl 蛋白酶 K（25 mg/ml）和 10 μl 20% SDS。

步骤 5A 和 6A 终止 DNase I 消化反应，应迅速进行。

7A. 将样品于 37℃孵育过夜。

MNase 消化

1B. 用 1.3 ml MNase 消化缓冲液洗涤沉淀。

2B. 4℃，1000 r/min 离心 10 min，沉淀细胞核。小心弃上清，用 500 μl 含 1 mmol/L $CaCl_2$ 的 MNase 消化缓冲液重悬细胞核。

每 1 ml MNase 消化缓冲液加 1 μl 1 mol/L $CaCl_2$。

3B. 转移 100 μl 细胞核至含有 MNase 的微量离心管中。

MNase 应根据活性单位分成小份保存于－20℃。作为合宜的起始点，使用 0.5～50U 的滴定范围。必须根据 LM-PCR 结果优化 MNase 用量。

4B. 将样品于室温下孵育 5 min。

5B. 每个样品中加入 80 μl MNase 消化缓冲液和 20 μl MNase 终止缓冲液。

6B. 加入 3 μl 蛋白酶 K（25mg/ml）和 10 μl 20% SDS。

步骤 5B 和 6B 终止 MNase 消化，应迅速进行。

7B. 37℃过夜孵育样品。

限制性内切核酸酶消化

1C. 用 1.3 ml 限制性内切核酸酶消化缓冲液洗涤细胞核。

2C. 4℃，1000 r/min 离心 10 min，沉淀细胞核。小心弃上清。用与酶一起提供的 250 μl 最适限制性内切核酸酶消化缓冲液（1×）重悬细胞核。

为了在该步中使用，将购买了限制性内切核酸酶的公司所提供或建议的 10×缓冲液进行稀释。如果分析多种限制性内切核酸酶，应使用相容的缓冲液，或将步骤 1C 中的样品分成多个等份。

3C. 将 50 μl 细胞核转移至含有限制性内切核酸酶的微量离心管中。

首先试用 2 μl 浓缩的限制性内切核酸酶（10 U/μl）进行消化是合宜的。目的是加入足以在随后的 LM-PCR 分析中检测到产物的酶量。但是，在控制区无活性的细胞中，这种核切割产生的 LM-PCR 条带应在曝光强度上远不及随后的体外对照切割所产生的条带（见下面的步骤 5）。如果核切割效率太高，就应减小酶的用量。

4C. 37℃孵育样品 10 min。

为了确保极限限制性内切核酸酶消化，可能需要优化消化的时间和酶的用量。

5C. 加入 50 μl 2×蛋白酶 K 缓冲液，55℃孵育 1 h。

该步终止限制性内切核酸酶消化。

6C. 加入 50 μl 2×蛋白酶 K 缓冲液，50 μl 限制性内切核酸酶消化缓冲液和 3 μl 蛋白酶 K（25 mg/ml）。

7C. 37℃孵育过夜。

步骤 4：纯化 DNA

1. 用 200 μl 酚：氯仿（pH 8）抽提样品。

为了防止剪切基因组 DNA，不要涡旋样品，取而代之的是通过轻弹管壁或振荡混匀，以充分混合水相和酚：氯仿层。

2. 在微量离心机中高速离心样品 5 min，小心转移水相层至新的离心管中。

基因组 DNA 可能表现“线”状。缓慢吸出水相层，保持“线”状 DNA，但尽量减少扰动酚：氯仿界面。

3. 用 200 μl 氯仿抽提样品。采取与步骤 1 相同的预防措施。

4. 在微量离心机中高速离心样品 5 min，小心转移水相层至新的微量离心管。采取与步骤 2 相同的预防措施。

5. 每个样品加入 2 μl 热处理的 RNase A（10 mg/ml）。37℃孵育 2 h。

6. 用 200 μl 酚：氯仿（pH 8）抽提样品。采取与步骤 1 相同的预防措施。

7. 在微量离心机中高速离心样品 5 min，小心转移水相层至新的微量离心管。采取与步骤 2 相同的预防措施。

在此阶段，界面通常不清晰。避免吸取界面，如果吸进少量界面时，在氯仿抽提中去除。

8. 用 200 μl 氯仿抽提样品。采取与步骤 1 相同的预防措施。

9. 在离心机中高速离心样品 5 min，小心转移水相层至新的离心管。

在此阶段，不要吸取任何云雾状界面，最好留一些水相层。

10. 通过加入 1∶10 体积的 3 mol/L 乙酸钠（pH 5.2）和 2.5 体积的冰冷乙醇，沉淀消化的基因组 DNA。样品于－20℃孵育约 1 h。

11. 在微量离心机中，4℃下高速离心 10 min，以沉淀 DNA。

12. 弃上清。用冰冷的 70％乙醇洗涤 DNA 沉淀。4℃下高速离心样品 1～2 min。

13. 弃上清。在操作台面上或用无热 SpeedVac 干燥 DNA。

小心不要使 DNA 过于干燥，否则难以重新溶解。

14. 重悬 DNA 于 50 μl H_2O 或者 TE（10 mmol/L Tris-HCl，pH 7.5，1 mmol/L EDTA）中。

轻轻地重悬 DNA；不要涡旋或剧烈吹打。可能有必要使 DNA 在 4℃下溶解过夜。如果溶液变得很黏稠，可以用 25 μl H_2O 稀释。

15. 对于 MNase 处理的 DNA，重复步骤 5～14。

在激酶反应前彻底去除 RNA 很关键。可能需要在琼脂糖凝胶上检查 DNA，以确定 RNA 是否仍然存在。

步骤 5：制备用于 LM-PCR 的 DNA

根据体内消化的不同，用于 LM-PCR 的 DNA 的制备亦不同：（A）DNase I；（B）MNase；（C）限制性内切核酸酶。

DNase I 消化 DNA

1A. 在第一天，准备好用于 LM-PCR 的 DNA。用分光光度计确定 DNA 浓度（$OD_{260/280}$）。取 1 μg DNA 用于 LM-PCR。

MNase 消化 DNA

1B. 用分光光度计确定 DNA 浓度。在磷酸化反应中用 1 μg DNA。

在无核小体的 DNA 和核小体连接 DNA 中，MNase 消化产生平末端切割。在切割反应中除去了 5′磷酸基团。为了进行随后的单向接头连接，必须恢复磷酸基。

2B. 在含有 1 mmol/L ATP 的 50 μl 总体积中进行磷酸化反应。

如果随连接酶提供的缓冲液与 T4 多核苷酸激酶兼容，就可以在连接缓冲液中进行磷酸化反应。这些缓冲液通常已含有 ATP，而由公司提供的激酶缓冲液缺乏 ATP。因此，磷酸化反应含有 5 μl 10×连接缓冲液、1 μg MNase 消化 DNA、H_2O 和 1.5 μl T4 多核苷酸激酶（New England Biolabs，10 U/μl）。

3B. 将反应在 37℃下孵育 1 h。

4B. 加 150 μl H_2O 至混合物中，并用 200 μl 酚∶氯仿（pH 8）抽提。

不要涡旋 DNA 样品，轻轻地、彻底地混匀。

5B. 在微量离心机中高速离心样品 5 min。将水相转移到新的微量离心管。

6B. 用 200 μl 氯仿抽提样品。

不要涡旋 DNA 样品，轻轻地、彻底地混匀。

7B. 在微量离心机中高速离心样品 5 min，将水相转移到新的微量离心管。

8B. 用 1/10 体积的乙酸钠（pH 5.2）和 2.5 体积的冰冷乙醇沉淀 DNA，将样品置于干冰上 10 min。

9B. 在微量离心机中 4℃下高速离心 10 min，以沉淀 DNA，弃上清。用 70%乙醇洗涤 DNA 沉淀。

10B. 在微量离心机中高速离心样品 1～2 min。小心移除上清。在操作台上或用无热 SpeedVac 干燥 DNA。

小心不要使 DNA 过于干燥。

11B. 将 DNA 重悬于 5 μl H_2O 或 TE（10 mmol/L Tris-HCl，pH 7.5，1 mmol/L EDTA）中，直接进行连接反应。

要检测无蛋白质区的 MNase 消化产物，不进行第一链合成。MNase 能切刻核小体内的 DNA。第一链合成检测这些切刻，从而产生混乱的结果。如上文所述，MNase 在无蛋白质区产生平末端。因此，在此阶段单向接头的连接将使这些切割得以检测。

限制性内切核酸酶消化 DNA

1C. 用分光光度计确定 DNA 浓度（$OD_{260/280}$）。取 1 μg DNA 进行体外限制性内切核酸酶消化。

体外限制性内切核酸酶消化是为了 PCR 扩增后输入 DNA（input DNA）的归一化。对体外消化而言，所选择的限制性内切核酸酶切割位点相对于用于 LM-PCR 的三个引物，应位于在体内限制性内切核酸酶位点的上游。

2C. 在 20 μl 的总体积中进行限制性内切核酸酶消化，消化反应在 37℃下孵育约 2 h。

将体外限制性内切核酸酶消化进行到彻底。

3C. 将限制性内切核酸酶在 65℃下热灭活 20 min。

4C. 取 5 μl 消化反应产物（0.25 μg DNA）用于 LM-PCR 第一链合成。

步骤 6：LM-PCR 第 1 天

配制 20 μmol/L 单向接头溶液

1. 凝胶纯化或 HPLC 纯化 LM-PCR 1 和 LM-PCR 2 寡核苷酸。

同时凝胶纯化用于 LM-PCR 的所有寡核苷酸是合宜的，也应该凝胶纯化引物 1～3。LM-PCR 1 和 LM-PCR 2 寡核苷酸长度不同，结果它们的退火形成的接头只有一端为平末端，这使得该接头可以与第一链合成步骤的平末端产物单向连接。

2. 配制 20 μmol/L 单向接头混合液：20 μmol/L LM-PCR 1、20 μmol/L LM-PCR 2、250 mmol/L Tris-HCl（pH 7.7）。

该混合液可配置成大体积，然后分成小等份；保存于－20℃，备用。

3. 将接头混合液于 95℃变性 5 min。

4. 简单离心混合液将管壁上的冷凝液滴收集到管内溶液中。将 20 μmol/L 单向接头混合液置于 70℃加热块中，将加热块从热源移开，并让其缓慢冷却至室温（大致需要 1 h）。

5. 将接头混合液在加热块中于 4℃孵育过夜。

6. 分装成小等份，－20℃保存。

接头混合液始终置于冰上解冻。

末端标记引物 3

1. 用［γ-^{32}P］ATP（6000 Ci/mmol）磷酸化引物 3。激酶反应含有以下组分：4.0 μl 10×多核苷酸激酶缓冲液、7.0 μl（70pmol）引物 3、4.0 μl［γ-^{32}P］ATP、23.5 μl H_2O 和 1.5 μl T4 多核苷酸激酶（New England Biolabs，10 U/μl）。

2. 在 37℃下孵育反应混合物 30～60 min。

3. 从游离的［γ-^{32}P］ATP 中纯化放射性标记的引物。

这可以用 Stratagene 推式柱（见方案 3.1）或相似的方法完成。

第一链合成反应

不要以 MNase 处理的 DNA 进行第一链合成反应，而是将模拟（mock）第一链混合液、稀释混合液和连接混合液合并在一起用于立即连接。第一链合成使目的 DNA 序列中的单链切口得以检测。单轮 PCR 将使基因特异性引物周围序列上的单链切口转化为平末端双链 DNA 片段，这些 DNA 片段随后可参与到与单向接头的连接反应中。

1. 往 PCR 管中加入 1～2 μg 的体内消化 DNA，使其总体积为 5 μl。将样品置于冰上。

加入到第一链合成反应中的 DNA 量为 0.25 μg（5μl），该 DNA 来源于限制性内切核酸酶消化。

2. 在冰上准备第一链合成反应混合液：（每个反应）6.0 μl 5×第一链缓冲液、0.3 μl（0.3 pmol）引物 1、0.25 μl 25 mmol/L dNTP 混合液和 18.2 μl H_2O。使用前，立即加入 0.25 μl pfu 聚合酶，混匀并简单离心。

用于 LM-PCR 的聚合酶是关键。Garrity 和 Wold（1992）进行了不同 DNA 聚合酶的比较，在该研究中没有检验 DNA 聚合酶 pfu，但我们已发现它对于 LM-PCR 很奏效。本方案用 pfu 作为 DNA 聚合酶。

3. 每个样品中加入 25 μl 冰冷的第一链合成反应混合液。

将样品保持在冰上，以防止 DNA 聚合酶的假活性。

4. 将样品置于自动热循环仪中，并按照以下条件进行第一链合成：95℃ 5 min（变性），60℃ 30 min（退火），76℃ 10 min（延伸）。

对于 MNase 样品，配制模拟第一链缓冲液（每个反应 25 μl），每个反应加入 19.0 μl H_2O 和 6.0 μl 5×第一链缓冲液。在冰上每个样品中加入 25 μl，并按以下描述继续进行。

5. 在冰上准备稀释液（每个反应 20 μl），每个反应加入 2.2 μl 1 mol/L Tris-HCl（pH 7.5）、0.35 μl 1 mol/L $MgCl_2$、2.0 μl 0.5 mol/L 二硫苏糖醇（DTT）、0.25 μl 无 DNase 的牛血清白蛋白（BSA）（10 mg/ml）和 15.2 μl H_2O。

6. 在冰上准备连接液（每个反应 25 μl），每个反应加入 0.25 μl 1 mol/L $MgCl_2$、1.0 μl 0.5 mol/L DTT、0.75 μl 100 mmol/L rATP、0.13 μl BSA（10 mg/ml）和 16.87 μl H_2O。每个反应使用前，立即加入 5.0 μl 20 μmol/L 单向接头和 1.0 μl（3U）T4 DNA 连接酶。

7. 在每个样品中加入 20 μl 冰冷的稀释液。

8. 在每个样品中加入 25 μl 冰冷的连接液。

9. 17℃下孵育样品过夜（大于 12 h)。

步骤 7：LM-PCR 第 2 天

1. 在冰上配制沉淀混合液：加入（每个反应）8.4 μl 3 mol/L 乙酸钠（pH 7）和 1.0 μl 酵母 tRNA（10 mg/ml)。

2. 将连接样品转移至新的离心管。每个样品中加入 9.4 μl 沉淀混合液和 220 μl 冰冷乙醇。在干冰上孵育 10 min 或于－20℃孵育 2 h。

3. 在微量离心机中 4℃下高速离心 10 min，沉淀 DNA。

4. 弃上清，用冰冷的 70％乙醇洗涤 DNA 沉淀，4℃下高速离心样品 1～2 min。

5. 弃上清并在操作台上或用无热 SpeedVac 干燥 DNA。

小心不要使 DNA 过于干燥。

6. 重悬 DNA 于 70 μl H_2O 中，将样品置于室温下直到 DNA 溶解（15～30 min)。

7. 在冰上配制扩增混合液（每个反应 30 μl)。每个反应加入 20 μl 5×扩增缓冲液、1.0 μl（10 pmol）LM-PCR 1、1.0 μl（10 pmol）引物 2、0.8 μl 25 mmol/L dNTP 混合液和 7.2 μl H_2O。

8. 将 30 μl 扩增混合液和 70 μl 样品合并在一个放在冰上的新 PCR 管中。

9. 在冰上配制 pfu 稀释混合液。每个反应加入 0.2 μl 5×扩增缓冲液、0.63 μl H_2O 和 0.17 μl pfu。给冰上的每个样品中加入 1.0 μl pfu 稀释混合液，混匀样品并简单离心。

pfu 聚合酶的用量可能需要滴定，因为太多或太少都会出现问题。

10. 如下进行 PCR：DNase I（21 个循环)、MNase（22～23 个循环)、限制性内切核酸酶（18 个循环)。使用下列条件：第一个循环，95℃变性 4 min，其余循环 95℃变性 1 min。退火步骤在引物 2 的最适杂交温度下进行 2 min。第一个循环 76℃延伸 5 min，以后的每一个循环延伸 15 s。最后一次延伸后，76℃下再进行 5 min 延伸。

理想的情况下，引物 2 的退火温度应比引物 2 计算出的熔解温度（T_m）高 2～5℃（或者高于接头引物的 T_m，接头引物 T_m 较低)，但是最适退火温度必须依据经验确定。我们以 63～64℃的退火温度对于 G/C 含量为 56％的引物（25 mer）效果好，但是对于每一个引物此步都必须优化。

11. 在冰上配制标记混合液（每个反应 4.5 μl)：加入（每个反应）1.0 μl 5×扩增缓冲液、2.0 μl（2.0pmol）末端标记的引物 3、0.4 μl 25 mmol/L dNTP 混合液和 0.8 μl H_2O。使用前，每个反应立即加入 0.3 μl pfu。

12. 在冰上，每个样品中加入 4.5 μl 标记混合液。进行标记 PCR 的三个循环，对于第一个循环，95℃变性 4 min；对于第 2 和第 3 个循环，95℃变性 1 min；在引物 3 的最适退火温度下退火 2 min；76℃下延伸 10 min。

如果背景或放射性标记扩增产物的量过高，标记 PCR 可以进行两个循环。此外，引物 3 的退火温度必须高于引物 2 的退火温度，这将排除引物 2 在标记循环期间和模板退火。

13. 转移 LM-PCR 样品至新的微量离心管并于每个样品中加入 300 μl 终止液。

14. 用 400 μl 酚：氯仿（pH 8）抽提样品。在微量离心机中室温下离心 5 min。转移水相至新的微量离心管。

该步骤可以省略，但是如果凝胶泳道上普遍的背景放射性高，应进行此步。

15. 在新的微量离心管中加入 94 μl 样品和 235 μl 冰冷乙醇。置于干冰上 10 min 以沉淀 DNA。各样品其余部分作适当的遮挡保存于－20℃。

16. 在微量离心机中 4℃下高速离心样品 10 min。弃上清（放射性的）并用冰冷的 70%乙醇洗涤 DNA 沉淀。

17. 4℃下高速离心样品 1～2 min。弃上清（放射性的）并在操作台上或用无热 SpeedVac 干燥 DNA。

18. 重悬 DNA 沉淀于 6.0 μl 甲酰胺上样缓冲液中。

19. 通过煮沸 2 min 变性样品。简单离心收集管壁上的冷凝液滴。

用盖锁（Lid Loks）防止煮沸过程中管盖崩开。

20. 将各样品全部的 6.0 μl 上样到 DNA 测序胶（8%变性聚丙烯酰胺）的 6 mm 上样孔中。60 W 下电泳约 1.5 h。

1.5 h 后，放射性引物应位于凝胶的底部，因此，可以切去凝胶的底部（稍高于溴酚蓝染料），以防止这部分放射性模糊所期望的信号，应弃于放射性废物中。

21. 干燥凝胶并在胶片或磷屏成像仪上曝光。

体内 DMS 足迹

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

乙酸铵（8 mol/L）

细胞

dH_2O

硫酸二甲酯（DMS）〈!〉

DMS 终止缓冲液

- 1.5 mol/L 乙酸钠（pH 7.0）
- 1 mol/L β-巯基乙醇〈!〉
- 100 μg/ml 酵母 tRNA

乙醇（100% and 75%）〈!〉

基因组 DNA（纯化的，或用作对照的质粒基因组克隆）

异丙醇〈!〉

LM-PCR 试剂

这些试剂与上面步骤 1 使用的试剂相同（DNase I、MNase 和限制性内切核酸

酶足迹程序)。

裂解缓冲液

1 mmol/L Tris-HCl (pH 7.5)

400 mmol/L NaCl

2 mmol/L EDTA

0.2% SDS〈!〉

0.2 mg/ml 蛋白酶 K〈!〉

立即使用前配制。

酚:氯仿〈!〉

磷酸缓冲盐

哌啶〈!〉

乙酸钠(3 mol/L,pH 5.2 和 3 mol/L,pH 7.0)

TE

10 mmol/L Tris-HCl (pH 7.5)

2 mmol/L EDTA

器材

离心机

锥形管(50 ml)

干冰

化学通风柜

90℃加热块

培养箱

LM-PCR 仪

微量离心管

巴斯德吸液管

移液器或真空吸液器

聚丙烯管(15 ml)

摇床

SpeedVac 浓缩器

涡旋器

水浴,设置在 37℃

程序

步骤 1:配制缓冲液

配制缓冲液使用的试剂与上面步骤 1 所用的 LM-PCR 试剂相同(DNase I、MNase 和限制性内切核酸酶足迹程序)。

步骤 2：用 DMS 处理细胞

对于贴壁细胞（15 cm 平板）

1. 37℃下预热培养基和 PBS。

用于 DMS 处理的培养基和细胞培养基相同。

2. 从细胞中移除培养基。

3. 在化学通风柜中，每 1 ml 培养基中加入 1 μl DMS，从而使其终浓度为 0.1%。混匀并立即加入到细胞中。使用足量的溶液，以轻易覆盖细胞，室温下孵育 2 min。

DMS 溶液应在使用前新鲜配制，可能需要对 DMS 浓度进行滴定，以确保最佳的反应。

4. 用一次性移液管移除含有 DMS 的培养基并用 PBS 洗涤细胞。

当优化 DMS 足迹时，DMS 处理的时间可有所变动，确保液态和固态 DMS 废物的恰当处理。

5. 用一次性移液管或真空吸液器（在化学通风柜中）移除 PBS 并用 PBS 重复洗涤三次。

每次洗涤期间细胞轻摇 30 s。最后一次洗涤后，可以从化学通风柜中取出细胞。

6. 最后一次 PBS 洗涤后，用 1.5 ml 裂解缓冲液裂解细胞。小心轻摇细胞使溶液覆盖细胞。室温下再孵育 5 min。

7. 通过刮取从平板上移除细胞裂解物并且转移至 15 ml 的一次性聚丙烯管中。

8. 37℃下孵育细胞裂解物 3～5 h，每隔一定时间进行混匀。

对于悬浮细胞

1. 对于每个样品，在两个一次性 50 ml 锥形管中加入等份的 49 ml 冰冷 PBS。在化学通风柜中放置 37℃水浴锅或加热块用于 DMS 处理，为每个样品配制 2.7 ml 裂解液。

2. 将悬浮细胞（约 1×10^8 个细胞）置于 50 ml 锥形管中，并通过 1500 r/min 离心 5 min 沉淀细胞。

3. 重悬细胞沉淀于 1 ml 培养基中（预热至 37℃），并转移至 1.5 ml 离心管，置于 37℃水浴锅或加热块中。

4. 在化学通风柜中，通过在 100%乙醇中溶解 DMS 配制 10%的 DMS 溶液，通过涡旋混匀溶液然后简单离心。

该溶液应于使用前新鲜配制。

5. 每个样品中加入 10 μl 10% DMS 溶液。通过颠倒密封的管混匀，并在 37℃孵育 1 min。

可以改变 DMS 的用量和处理时间，以优化反应和 DMS 足迹。

6. 转移 DMS 处理的样品至等份的 49 ml 冰冷的 PBS 中并混匀。1500 r/min 离心 5 min 沉淀细胞，移除上清并恰当处理废物。

7. 重悬细胞沉淀于 1 ml 冰冷的 PBS 中并用第二等份的 49 ml PBS 重复步骤 6。

8. 重悬细胞沉淀于 300 μl 冰冷的 PBS 中。每份样品置于 15 ml 的聚丙烯管中，并加入 2.7 ml 裂解缓冲液。轻轻地、彻底地混匀样品。

9. 37℃下孵育细胞裂解物 3～5 h，每隔一定时间通过颠倒混匀。

步骤 3：纯化 DNA

1. 用等体积的酚∶氯仿抽提样品。通过轻轻地颠倒试管数次将样品彻底混匀。在台式离心机中 3000 r/min 离心样品 10 min。

为了避免剪切基因组 DNA，不要涡旋。

2. 移除酚∶氯仿层（下层）并弃于酚废物中。重复水相层的酚∶氯仿抽提。

通过将巴斯德吸管穿过水相层进入酚∶氯仿层小心吸取以移除有机相，这可能比移除水相层至新管更容易，因为基因组 DNA 将是黏稠的，这种方法也将减少对基因组 DNA 的剪切。

3. 移除酚∶氯仿层并弃于酚废物中，用等体积的氯仿抽提样品两次。

4. 用等体积的异丙醇沉淀 DNA。

将看到白色、线状的 DNA 沉淀。

5. 用密封的巴斯德吸管通过缠绕收集 DNA。

对于 DNA 缠绕的详细描述，见 Ausubel 等（1994，单元 15.5）。

6. 在 3 ml TE（10 mmol/L Tris-HCl，pH 7.5，1 mmol/L EDTA）中轻轻摇动缠绕的 DNA，使 DNA 重悬。

为了使 DNA 重回溶液，可能需花费数小时至过夜，不要涡旋。

7. 用 1/10 体积的 3 mol/L 乙酸钠（pH 5.2）和 2 倍体积的乙醇沉淀 DNA，通过缠绕收集 DNA。

8. 在 4℃下将缠绕的 DNA 重悬于 200～500 μl TE（pH 7.5）中。

为了使 DNA 重回溶液，可能需花费数小时至过夜，不要涡旋。

9. 用分光光度计确定 DNA 浓度，并将终浓度调至 1～1.5 mg/ml。

该制剂中大部分是 RNA，RNA 在哌啶切割步骤中将会被消除，4℃保存样品。

步骤 4：用哌啶处理 DNA

1. 用 50 μl DMS 终止液和 750 μl 乙醇沉淀 200 μl DNA。将样品在干冰上放置 30 min。将样品在微量离心机中 4℃下离心 10 min。

2. 用 75%乙醇洗涤 DNA 沉淀。4℃下离心样品 1～2 min 并移除上清。

在此阶段，不要使 DNA 沉淀完全干燥。

3. 通过将哌啶按 1∶10 体积加入到水中配制稀释的哌啶储液（1 mol/L）。每个样品中加入 200 μl 稀释的哌啶并通过涡旋重悬。

可能需花费 15～30 min 使 DNA 溶解于哌啶中，期间每隔一定时间涡旋。

4. 在微量离心机中简单离心样品，并在管上放置盖锁以防止加热过程中管盖崩出。将管置于放置在化学通风柜中的 90℃加热块中，孵育 30 min。

本步骤将在甲基化的鸟嘌呤处切割 DNA。

5. 在微量离心机中简单离心样品，收集管壁上的冷凝液滴，将样品在干冰上放置 10 min。

6. 通过在 SpeedVac 浓缩器中室温蒸发 1～2 h，除去哌啶。

7. 将 DNA 重悬于 360 μl TE（pH 7.5）中。用 1∶10 体积的 3 mol/L 乙酸钠（pH 7.0）和 2.5 倍体积的乙醇沉淀，将样品置于干冰上 10 min。

8. 在微量离心机中 4℃下高速离心样品 15 min。

9. 弃上清，将沉淀重悬于 500 μl TE（pH 7.5）中，用 170 μl 8 mol/L 乙酸钠和 670 μl 异丙醇沉淀 DNA。－20℃下孵育 2 h。

10. 在微量离心机中 4℃下高速离心样品 15 min。

11. 弃上清，用冰冷的 75％乙醇洗涤 DNA 沉淀。在微量离心机中高速离心样品1～2 min。

12. 弃上清，将 DNA 沉淀重悬于 50μl dH_2O 中。

13. 在产热 SpeedVac 浓缩器中蒸发 1 h，除去 dH_2O。

14. 将 DNA 沉淀重悬于 TE（pH 7.5），使其终浓度约为 1 μg/ml。

15. 室温下在微量离心机中高速离心样品 10 min。移除上清至新管，留下清晰的沉淀。

步骤 5：制备用于 LM-PCR 的 DNA

利用分光光度计确定 DNA 浓度，取 1 μg DNA 用于 LM-PCR。

步骤 6 和步骤 7：LM-PCR 第 1 天和 LM-PCR 第 2 天

按上述步骤 6 和步骤 7（DNase I、MNase 和限制性内切核酸酶足迹程序）进行。

附加说明

1. 为了解释体内 DNase I、MNase 和 DMS 的数据，必须做对照，使纯化的 DNA 进行同样的程序。可以用纯化的基因组 DNA 或质粒基因组克隆。MNase 和 DNase I 具有明显的序列偏好，因此，纯化的 DNA 对照对于受保护和高敏感核苷酸的鉴定是关键的。对于 DNase I 足迹，纯化的 DNA 对照可能有些难以解释，因为除了序列特异性 DNA 结合蛋白外，还有组蛋白可以改变 DNase I 的消化模式。如果体内足迹研究的目的是转录因子结合分析，那么基因的转录沉寂和活跃状态的比较可能比与纯化 DNA 的比较更为有用。

2. 解释体内足迹数据时，考虑的一个重要问题是被研究的细胞群体的性质。特别是细胞群体中表达某基因的细胞所占的比例可显著地影响结果的质量和有效性。如果仅有 25％的细胞是转录活跃的，将难以检测 DNA 结合蛋白对控制元件的占据。

3. 重要的是要牢记体内足迹技术不像标准的体外足迹技术，需要 PCR 扩增。PCR 扩增会提高或降低所观察到的保护或高敏感性的程度。如果没有观察到相互作用，改变扩增循环数可能是有利的。

4. PCR 扩增的需要大大增加了对体内足迹数据客观评价的需要，以及每项实验的多次重复以表明重复性的需要。在用一套特定引物的大多数单独的实验中，将观察到似乎受保护的或高敏感的几个核苷酸。然而，这些变化中只有部分变化在实验间将是可重

复的，并达到足够得出有力结论的程度。

5. 重要的是牢记 LM-PCR 产物中含有 25 bp 的额外序列，它是由连接和扩增的接头序列贡献的。在涉及确定那些位点位置的计算中，在受保护和高敏感的位点与 DNA 测序标准物相比较时，必须考虑这 25 bp。

疑难解答

高背景

可能原因：引物的退火温度可能不够严紧。

解决办法：对于引物 2 和引物 3，应该测试不同的退火温度。在标记反应期间，使用最高的可能温度以减少非特异性退火并增强引物 2 的排斥性是非常重要的。

可能原因：PCR 过度扩增。

解决办法：如果泳道过度曝光，可能需要减少 PCR 循环数。如果纯化的 DNA 对照和体内消化的 DNA 相比似乎过度扩增，则应减少 PCR 中模板的用量。

空白泳道

可能原因：对于 MNase 程序，MNase 处理 DNA 的激酶催化可能是低效的。

解决办法：在有大量 RNA 的情况下，会降低激酶化的效率。在琼脂糖凝胶上电泳 MNase 处理的样品以确定 RNA 是否存在。如果存在 RNA，则重复 RNase A 消化。另外，激酶化大量 DNA 将会降低反应的总体效率。

可能原因：体内消化水平太高或太低。

解决办法：对酶进行更广泛的滴定将会增加找到合适范围的可能性。在琼脂糖凝胶上可以分析消化的 DNA，以确定消化的水平。

可能原因：试剂质量差。

解决办法：精确配置试剂并使用高质量的组分是极其重要的。

可能原因：退火温度太高或引物组合不匹配。

解决办法：首先尝试调整引物 2 和引物 3 的退火温度。如果这样没有奏效，则引物可能不匹配，需要设计新的引物。

方案 9.3 染色质免疫沉淀

Gilmour 和 Lis（1984、1985、1986）发展了第一个染色质免疫沉淀（ChIP）分析，并将其作为监测大肠杆菌（*E. coli*）和果蝇（*Drosophila*）中 RNA 聚合酶 II 与转录基因和蓄势基因结合的一种技术。在早期的研究中，用来自透射仪的紫外线将蛋白质不可逆地交联到 DNA 上，然后，交联的染色质用超声波打断或限制性内切核酸酶切割，产生较小的 DNA 片段，接着与期望的抗体免疫沉淀。之后，用蛋白酶处理沉淀的蛋白质-DNA加合物，纯化 DNA，并通过斑点印迹或 Southern 印迹用来源于克隆的目的 DNA 片段的放射性标记探针分析 DNA。此外，切刻平移放射性标记沉淀的 DNA 片

段，并将其用于探测 Southern 印迹膜（该膜上含有对应于目的区域的未标记的质粒DNA）。

甲醛作为蛋白质-DNA 和蛋白质-蛋白质的可逆交联试剂用于 ChIP，最早是由 Solomon 和 Varshavsky 开创的（Solomon et al. 1988）。在他们原创性的研究中，监测了组蛋白 H4 和 RNA 聚合酶 II 与果蝇 *hsp70* 基因的结合。该研究中，首先将细胞在热激前后用甲醛处理，接着进行 DNA 的剪切或限制性内切核酸酶消化。之后，用特定抗体对包含 H4 组蛋白或 RNA 聚合酶 II 的蛋白质-DNA 交联复合物进行免疫沉淀，再通过加热逆转免疫沉淀物中的共价交联，纯化免疫沉淀物中的 DNA；最后通过狭缝印迹或 Southern 印迹分析纯化的 DNA 片段。

将 PCR 加入 ChIP 分析，用于检测沉淀的 DNA 片段，是将其作为现代 ChIP 程序（见图 8.1）的最后环节。Hecht 和 Grunstein（1996）首先将该检测方法用于他们对酿酒酵母（*S. cerevisiae*）中 SIR 蛋白相互作用的研究。此后不久，Rundlett 等（1998）使用带有 PCR 检测步骤的 ChIP 方法，分析特定基因座上的组蛋白修饰。大约与此同时，对 ChIP 程序进行了修改，使其用于哺乳动物细胞，在该程序中，首先用紫外线交联，随后用甲醛交联（Boyd and Farnham 1997，1998；Wathelet et al. 1998；Parekh and Maniatis 1999）。重要的是要注意，尽管作为一种普遍而又强大的技术，ChIP 的出现恰逢将甲醛作为交联试剂使用，但是 ChIP 也可以在没有交联的情况下进行（即自然 ChIP 或 N-ChIP），用于检测在染色质处理和免疫沉淀期间与 DNA 保持稳定结合的蛋白质（O'Neill and Turner 2003）。

下面描述的用于哺乳动物细胞的 ChIP 方案代表一个标准的 ChIP 程序，该程序来源于上面引用的方案。已经使用了这一基本方案的许多变通方案，可以在文献（Kuo and Allis 1999；Orlando 2000；Johnson and Bresnick 2002；O'Neill and Turner 2003；Ren and Dynlacht 2004）或众多的网站上，包括提供的 ChIP 抗体和试剂的公司的网站上发现这些变通方案。特别是 Abcam 网站，它包含详细的 ChIP 方案，以及关于实验设计和疑难解答的非常好的提示。描述 ChIP 用于全基因组研究方法的论文也已发表（Bernstein et al. 2004；Ciccone et al. 2004；Ren and Dynlacht 2004；Elnitski et al. 2006）。此外，以异常少的细胞开始 ChIP 研究的专门实验方案也已经发表（O'Neill et al. 2006；Acevedo et al. 2007；Attema et al. 2007）。

时线与安排

除了下面列出的那些试剂外，仅本程序所要求的特殊试剂包括待分析的细胞、特异性识别自然目的蛋白（或翻译后修饰型蛋白）的抗体，以及 PCR 引物或实时 PCR 引物，这些引物用于分析经沉淀和解交联的 DNA。在第一天，通过向生长的细胞中加入甲醛进行交联，然后准备染色质，通过超声波剪切（或在其他的程序中，用 MNase 进行酶消化），预清除以减少非特异性免疫沉淀。免疫沉淀一般以特定抗体过夜进行。在第二天，加入蛋白 A-琼脂糖或者蛋白 G-琼脂糖，孵育后，通过离心和洗涤分离免疫沉淀。然后，从树脂上洗脱蛋白质-DNA 复合物，并加热样品（通常过夜）以逆转共价交联。最后，纯化 DNA 片段并通过 PCR 或实时 PCR 进行分析。

提纲

步骤 1：准备缓冲液（2 h）。
步骤 2：交联和染色质制备（2 h）。
步骤 3：染色质预清除/免疫沉淀（2 h 和过夜）。
步骤 4：免疫沉淀洗涤和洗脱/解交联（2 h 和过夜）。
步骤 5：纯化 DNA（1 h）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

琼脂糖凝胶
自然目的蛋白抗体
抗体（对照）
细胞裂解液
　5 mmol/L PIPES（pH 8.0）
　85 mmol/L KCl〈!〉
　0.5% NP-40（Nonidet P-40）〈!〉
　4℃保存。
细胞（每个实验 $5\times10^7\sim1\times10^8$ 个）
稀释缓冲液
　16.7 mmol/L Tris-HCl（pH 8.0）
　167 mmol/L NaCl
　1.2 mmol/L EDTA
　0.01% SDS〈!〉
　1.1% Triton X-100〈!〉
　4℃保存。
EDTA
洗脱缓冲液
　50 mmol/L Tris-HCl（pH 8.0）
　10 mmol/L EDTA
　1% SDS〈!〉
　室温保存。
甲醛（37%）〈!〉
甘氨酸（1.375 mol/L）
生长培养基
高盐洗涤缓冲液

50 mmol/L HEPES（pH 7.9）
500 mmol/L NaCl
1 mmol/L EDTA
0.1% SDS〈!〉
1% Triton X-100〈!〉
0.1% 脱氧胆酸〈!〉
4℃保存。

微球菌核酸酶（MNase）

MNase 消化缓冲液
10 mmol/L Tris-HCl（pH 7.4）
15 mmol/L NaCl
60 mmol/L KCl〈!〉
1 mmol/L $CaCl_2$〈!〉
0.15 mmol/L 精胺〈!〉
0.5 mmol/L 亚精胺〈!〉
4℃保存。

细胞核裂解缓冲液
50 mmol/L Tris-HCl（pH 8.0）
10 mmol/L EDTA
1% SDS〈!〉
室温保存。

PCR 或实时 PCR 引物
酚：氯仿〈!〉（pH 8.0）
1×磷酸缓冲盐（PBS）
蛋白酶抑制剂（Roche）
蛋白 A 或蛋白 G-琼脂糖/鲑鱼精 DNA 微珠（Millipore）
蛋白酶 K（20 μg/μl）〈!〉

TE
10 mmol/L Tris-HCl（pH 8.0）
1 mmol/L EDTA
4℃保存。

设备

离心机（台式）
锥形管（15 ml）
凝胶电泳仪
凝胶上样枪头（狭窄的，用于吸出）
冰

培养箱
微量离心机
微量离心管（1.5 ml）
PCR 仪
转子
超声波发生器
分光光度计
离心柱（QIAGEN）
组织培养皿或培养瓶

步骤 1：准备缓冲液

步骤 2：交联和染色质制备

每个实验条件始于 5×10^7～1×10^8 个细胞。

1. 每毫升细胞培养物（在组织培养板上或组织培养瓶中的生长培养基中）加 27 μl 37%的甲醛，同时在室温下慢慢摇晃 10 min。

甲醛的终浓度为 1%。

2. 每毫升培养液中加 100 μl 1.375 mol/L 的甘氨酸，终止交联。

3. 对于贴壁细胞，移除生长培养基并适当地处理。用 10 ml 冰冷的 1×PBS 洗涤细胞两次，然后刮取细胞，收集到 15 ml 锥形管内的 1×PBS 中。4℃下 1500 r/min 离心细胞 10 min。对于不贴壁的细胞，转移细胞至 15 ml 锥形管中，4℃下 1500 r/min 离心细胞 10 min，并用 10 ml 冰冷的 1×PBS 洗涤细胞两次。

4. 重悬细胞沉淀于 10 ml 细胞裂解缓冲液（冰冷）中，并在冰上孵育 10 min。

5. 在台式离心机中，4℃下 1000 r/min 离心 10 min，沉淀细胞核，小心吸取上清。

6. （可选）重悬细胞核于 500 μl MNase 消化缓冲液中，加入微球菌核酸酶并在冰上孵育 1 h。

步骤 6～7 是可以选择的。根据步骤 9 中最终染色质的大小将需要优化 MNase 的用量。50～300U 的范围就是一个合适的起点。

7. 加入 EDTA，至终浓度为 50 mmol/L，终止消化。

8. 重悬细胞核沉淀于 1 ml 补充蛋白酶抑制剂的细胞核裂解缓冲液中，并在冰上孵育 10 min。

9. 进行超声处理。小心确保样品不起泡沫，并保持尽可能的低温。

超声条件将需要优化。一般在 6.0 的输出功率下，6 个 15 s 的脉冲加 45 s 的间歇期，就能奏效。适当地超声处理有助于分析每个样品。对于该分析，按如下所述解交联并纯化小体积染色质，然后在琼脂糖凝胶上分析。理想情况下，DNA 大小应为 300～1000 bp。

10. 转移染色质到 1.5 ml 离心管中并在－80℃冷冻储存。

步骤 3：染色质预净化和免疫沉淀

1. 在冰上解冻染色质。在微量离心机中，4℃下高速离心样品 15 min，沉淀析出的 SDS。转移上清至新微量离心管中。

在这一步中，将染色质分成小等份储存于－80℃是有益的。

2. 用分光光度计测量染色质的 DNA 浓度（A_{260}）。用去除 SDS 的细胞核裂解缓冲液作为空白。

一般来说，染色质的浓度应该大于 750 ng/μl，A_{260}/A_{280} 比值应为 1.4～1.6。

3. 为待用的每个抗体分装 100 μg 染色质于离心管中。

务必有一个无关的抗体作为对照，如 α-谷胱甘肽 *S*-转移酶。

4. 用补充有蛋白酶抑制剂的稀释缓冲液稀释染色质至终体积为 300 μl。

5. 为了预净化染色质，向染色质中加入 50 μl 蛋白 A 或蛋白 G-琼脂糖/鲑鱼精 DNA 微珠（Millipore），4℃下旋转 1～2 h。

使用大口径枪头或者剪掉末端的常规 Pipetman 枪头转移微珠。

6. 4℃下 3000 r/min 离心染色质样品 5 min。转移上清至新的微量离心管中。

7. 将相关的抗体加入染色质样品中，4℃下旋转过夜。

关于抗体的加入量，参考供应商提供的信息表，一般 5～10 μg 就足够。

步骤 4：免疫沉淀洗涤和洗脱

1. 往冰上的染色质样品中加入 50 μl 蛋白 A 或 G-琼脂糖/鲑鱼精 DNA（取决于抗体）微珠，4℃下旋转 2 h。

往样品中加入前，确保彻底重悬微珠浆液。

2. 4℃，3000 r/min 离心样品 2 min。

3. 从对照抗体样品的上清中收集 15 μl 作为 5%输入（input），将此样品保持在冰上备用。

4. 小心吸取所有样品的上清。

小心不要吸取任何微珠。狭窄的凝胶上样枪头对于小心吸取非常有用。

5. 所有样品中加入 1 ml 高盐洗涤缓冲液，室温旋转 10 min。

6. 室温下 3000 r/min 离心样品 2 min。

7. 小心吸取上清并加入 1 ml 高盐洗涤缓冲液，室温下旋转 10 min。

8. 重复步骤 6 和步骤 7 两次，一共用高盐洗涤缓冲液洗涤 4 次。

9. 吸除上清并用 TE 洗涤两次，洗涤方法同上。

10. 将微珠和输入样品重悬于补充了 1 μl 蛋白酶 K（20 μg/μl）的 300 μl 洗脱缓冲液中，55℃下孵育样品 2 h。

不要通过吹打重悬微珠，取而代之的是轻弹或者涡旋样品。

11. 65℃下孵育过夜，解交联。

步骤 5：纯化 DNA

1. 室温下全速离心样品 5 min。

2. 将上清转移至新的微量离心管中。

3. 此时，可以用离心柱（QIAGEN）或者标准的酚：氯仿抽提（见 Sambrook and Russell 2001，8.9）纯化 DNA。如果进行酚：氯仿抽提，应该加入酵母 tRNA 或糖原载体。

4. 将 DNA 重悬于 50 μl H_2O 或者 TE 中，并通过 PCR 或实时 PCR 进行分析。

附加说明

1. 实时 PCR 分析通常是分析免疫沉淀样品中特定 DNA 片段的首选方法，结果通常以“输入百分数”的值表现出来，该值是通过用实时 PCR 定量加入 ChIP 反应中的目的 DNA 片段的丰度来计算的，相对于最终的免疫沉淀物中发现的 DNA 片段的丰度。

2. 作为 ChIP 实验关键的阴性对照，可用实时 PCR 来监测其他不期望用所用的特定抗体沉淀下来的基因组 DNA 片段。最好是检查一系列跨越的假定相互作用位点的实时 PCR 引物对。假定相互作用位点上的“输入百分数”峰信号会加强证据，表明该相互作用真正是该位点特异的。作为第二个关键的阴性对照，应该用一个无关抗体进行平行的 ChIP 实验，该抗体针对一个与染色质上的目的位点不结合的蛋白质。

3. 正如在本方案介绍中所提到的，也可以在缺乏交联的情况下进行 ChIP 实验。在有交联的情况下进行的 ChIP 有时被称为 X-ChIP，而在缺少交联的情况下进行的 ChIP 有时被称为 N-ChIP。有关 X-ChIP 和 N-ChIP 优缺点的精彩讨论，以及详细的 N-ChIP 方案，见 O'Neill 和 Turner（2003）。在 Abcam 网站上也可获得相似的方案。N-ChIP 的优点包括更高的分辨率、更有效地测试抗体免疫沉淀的效率，以及消除交联损害抗体识别其靶蛋白能力的顾虑。N-ChIP 的缺点包括该方法仅限于像组蛋白这种与 DNA 结合稳定性很强的蛋白质，以及在染色质处理过程中核小体可能会重排的可能性。

4. 正如在本章节中题为“监测核小体存在和定位的 LM-PCR、PCR 及 ChIP 策略”的部分及在那部分引用的论文（见 O'Neill and Turner 2003）中所描述的，用于 ChIP 的染色质可单独通过 MNase 消化制备，而不用超声波制备。MNase 消化是打碎用于 N-ChIP 的染色质的首选方法，然而用于 X-ChIP 的染色质，通常是用超声波打碎，有时与 MNase 消化结合。

5. ChIP 的一个主要局限性是需要高度特异性识别靶蛋白的抗体，以及即使在交联后也能有效沉淀蛋白的抗体。重要的是要意识到抗体可与 ChIP 实验中的其他染色质结合蛋白交叉反应，即使抗体在 Western 印迹上显得高度特异。检查这种可能性的理想对照是在缺少靶蛋白的细胞（最好是在编码靶蛋白的基因内含有空白突变的细胞）中进行平行 ChIP 实验。在第 9 章中及文献 Beima 等（2006）中也讨论了这个问题。

6. 在第 8 章或第 9 章正文中讨论的 ChIP 的其他局限性包括：与阴性对照相比，该分析通常产生低信号，从而导致不确定的结果；该分析方法有限的分辨率；蛋白质与基因组位点的结合并不一定表示功能相关性。

疑难解答

无特异性信号

可能原因：抗体不能有效地从交联的或自然的染色质样品中沉淀蛋白质。

解决方案：可以用 Western 印迹监测沉淀效率。如果免疫沉淀效率低，即使孵育过夜也是如此，则应加入更多的抗体。此外，可能有必要鉴定或生产另一种抗体。商业上可获得的大部分抗体产生自肽。尽管抗肽抗体有时以高亲和力结合它们的靶抗原，但很多却并非如此。

可能原因：目的蛋白不能有效地交联到染色质上。

解决方案：在甲醛交联之前通过加入蛋白质-蛋白质交联试剂有时可以提高交联效率（Kurdistani and Grunstein 2003）。然而，有些蛋白质将不会有效地交联，因为这些蛋白质与染色质的结合具有瞬时的特点。增加交联时间可能有益。

可能原因：在交联的染色质中，抗体识别的蛋白质抗原表位被屏蔽。

解决方案：对于单克隆抗体或肽抗体而言，抗原表位屏蔽可能问题更多，因为这些抗体制剂通常仅识别单个抗原表位。因此，用完整的天然蛋白、融合蛋白或蛋白质大片段产生的多克隆抗体是首选。

可能原因：超声波处理后的 DNA 片段可能太大或者太小。

解决方案：在琼脂糖凝胶上监测 DNA 片段的大小，并调整超声时间，以产生长度为 300 ～1000 bp 的片段。

可能原因：甲醛交联时间可能太短或太长。

解决方案：测试不同的交联时间。

可能原因：洗涤期间可能破坏了蛋白 A-抗体的相互作用。

解决方案：降低洗涤的严紧性或尝试 G 蛋白；G 蛋白对一些同型抗体和来源于某些物种的抗体具有较高的亲和力。

阴性对照抗体的高背景

可能原因：预清除不够充分。

解决方案：增加用蛋白 A-琼脂糖树脂预清除的时间或者增加用于预清除的微珠的数量。

可能原因：缓冲液的基因组 DNA 或质粒 DNA 污染。

解决方案：进行无染色质模板的对照 PCR，检测可能的污染。

可能原因：蛋白 A 树脂可能质量差。

解决方案：测试不同批次或供应商的树脂。

参考文献

Acevedo L. G., Iniguez A. L., Holster H. L., Zhang X., Green R., and Farnham P. J. 2007. Genome-scale ChIP-chip analysis using 10，000 human cells. *BioTechniques* 43: 791-797.

Attema J. L., Papathanasiou P., Forsberg E. C., Xu J., Smale S. T., and Weissman I. L. 2007. Epigenetic

characterization of hematopoietic Stem cell differentiation using miniChIP and bisulfite sequencing analysis. *Proc. Natl. Acad. Sci.* 104: 12371-12376.

Ausubel F. M., Brent R. E., Kingston E., Moore D. D., Seidman J. G., Smith J. A., and Struhl K. 1994. *Current protocols in molecular biology*. John Wiley and Sons, New York.

Barski A., Cuddapah S., Cui K., Rob T. Y., Schones D. E., Wang Z., Wei G., Chepelev I., and Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* 129: 823-837.

Beima K. M., Miazgowicz M. M., Lewis M. D., Yan P. S., Huang T. H., and Weinmann A. S. 2006. T-bet binding to newly identified target gene promoters is cell type-independent but results in variable context-dependent functional effects. *J. Biol. Chem.* 281: 11992-12000.

Bernstein B. E., Humphrey E. L., Liu C. L., and Schreiber S. L. 2004. The use of chromatin immunoprecipitation assays in genome-wide analyses of histone modifications. *Methods Enzymol.* 376: 349-360.

Blais A. and Dynlacht B. D. 2005. Constructing transcriptional regulatory networks. *Genes Dev.* 19: 1499-1511.

Bosisio D., Marazzi I., Agresti A., Shimizu N., Bianchi M. E., and Natoli G. 2006. A hyper-dynamic equilibrium between promoter-bound and nucleoplasmicdimers controls NF-κB-dependent gene activity. *EMBO J.* 25: 798-810.

Boyd K. E. and Farnham P. J. 1997. Myc versus USF: Discrimination at the cad gene is determined by core promoter elements. *Mol. Cell. Biol.* 17: 2529-2537.

Buck M. J., and Lieb J. D. 2004. ChIP-chip: Considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics* 83: 349-360.

Chan S. W., Henderson I. R., and Jacobsen S. E. 2005. Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nat. Rev. Genet.* 6: 351-360.

Chaya D. and Zaret K. S. 2004. Sequential chromatin immunoprecipitation from animal tissues. *Methods Enzymol.* 376: 36 1-372.

Chen X., Wang J., Woltring D., Gerondakis S., and Shannon M. F. 2005. Histone dynamics on the interleukin-2 gene in response to T-cell activation. *Mol. Cell Biol.* 25: 3209-3219.

Ciccone D. N., Morshead K. B., and Oettinger M. A. 2004. Chromatin immunoprecipitation in the analysis of large chromatin domains across murine antigen receptor loci. *Methods Enzymol.* 376: 334-348.

Clark S. J., Harrison J., Paul C. L., and Frommer M. 1994. High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* 22: 2990-2997.

Clark S. J., Statham A., Stirzaker C., Molloy P. L., and Frommer M. 2006. DNA methylation: Bisuiphite modification and analysis. *Nat. Protoc.* 1: 2353-2364.

Cokus S. J., Feng S., Zhang X., Chen Z., Merriman B., Haudenschild C. D., Pradhan S., Nelson S. F., Pellegrini M., and Jacobsen S. E. 2008. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 452: 215-219.

Costello J. F., Smiraglia D. J., and Plass C. 2002. Restriction landmark genome scanning. *Methods* 27: 144-149.

Das P. M., Ramachandran K., vanWert J., and Singal R. 2004. Chromatin immunoprecipitation assay. *BioTechniques* 37: 961-969.

de Laat W. and Grosveld F. 2007. Inter-chromosomal gene regulation in the mammalian cell nucleus. *Curr. Opin. Genet. Dev.* 17: 456-464.

de Laat W., Klaus P., Kooren J., Noordermeer D., Paistra R. J., Simonis M., Splinter E., and Grosveld F. 2008. Three-dimensional organization of gene expression in erythroid cells. *Curr. Top. Dev. Biol.* 82: 117-139.

Dekker J., Rippe K., Dekker M., and Kleckner N. 2002. Capturing chromosome conformation. *Science* 295: 1306-1311.

Dong J., Jimi E., Zhong H., Hayden M. S., and Ghosh S. 2008. Repression of gene expression by unphosphorylated NF-κB p65 through epigenetic mechanisms. *Genes Dev.* 22: 1159-1173.

Drissen R., Palstra R. J., Gillemans N., Splinter E., Grosveld F., Philipsen S., and de Laat W. 2004. The active spatial organization of the β-globin locus requires the transcription factor EKLF. *Genes Dev.* 18: 2485-2490.

Elnitski L., Jin V. X., Farnham P. J., and Jones S. J. 2006. Locating mammalian transcription factor binding sites: A survey of computational and experimental techniques. *Genome Res.* 16: 1455-1464.

Enver T., Brewer A. C., and Patient R. K. 1985. Simian virus 40-mediated *cis* induction of the *Xenopus* β-globin DNase I hypersensitive site. *Nature* 318: 680-683.

Fascher K. D., Schmitz J., and Horz W. 1993. Structural and functional requirements for the chromatin transition at the PHO5 promoter in *Saccharomyces cerevisiae* upon PHO5 activation. *J. Mol. Biol.* 231: 658-667.

Fouse S. D., Shen Y., Pellegrini M., Cole S., Meissner A., Van Neste L., Jaenisch R., and Fan G. 2008. Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* 2: 160-169.

Fragoso G., John S., Roberts M. S., and Hager G. L. 1995. Nucleosome positioning on the MMTV LTR results from the frequency-biased occupancy of multiple frames. *Genes Dev.* 9: 1933-1947.

Frommer M., McDonald L. E., Millar D. S., Collis C. M., Watt F., Grigg G. W., Molloy P. L., and Paul C. L. 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci.* 89: 1827-1831.

Garrity P. A. and Wold B. J. 1992. Effects of different DNA polymerases in ligation mediated PCR: Enhanced genomic sequencing and in vivo footprinting. *Proc. Natl. Acad. Sci.* 89: 1021-1025.

Garrity P. A., Chen D., Rothenberg E. V., and Wold B. J. 1994. Interleukin-2 transcription is regulated in vivo at the level of coordinated binding of both constitutive and regulated factors. *Mol. Cell. Biol.* 14: 2159-2169.

Garrity P. A., Wold B., and Mueller P. R. 1995. Ligation-mediated PCR. In *PCR 2: A practical approach* (ed. M. J. McPherson et al.), p. 309. IRL Press at Oxford University Press, New York.

Giardina C., Pérez-Riba M., and Lis J. T. 1992. Promoter melting and TFIID complexes on *Drosophila* genes in vivo. *Genes Dev.* 6: 2190-2200.

Gilmour D. S. and Lis J. T. 1984. Detecting protein-DNA interactions in vivo: Distribution of RNA polymerase on specific bacterial genes. *Proc. Natl. Acad. Sci.* 81: 4275-4279.

Gilmour D. S. and Lis J. T. 1985. In vivo interactions of RNA polymerase II with genes of *Drosophila melanogaster*. *Mol. Cell. Biol.* 5: 2009-2018.

Gilmour D. S. and Lis J. T. 1986. RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in *Drosophila melanogaster cells*. *Mol. Cell. Biol.* 6: 3984-3989.

Gimelbrant A., Hutchinson J. N., Thompson B. R., and Chess A. 2007. Widespread monoallelic expression on human autosomes. *Science* 318: 1136-1140.

Göndör A., Rougier C., and Ohlsson R. 2008. High-resolution circular chromosome conformation capture assay. *Nat. Protoc.* 3: 303-313.

Gottschling D. E. 1992. Telomere-proximal DNA in Saccharomyces cerevisiae is refractory to methyltransferase activity in vivo. *Proc. Natl. Acad. Sci.* 89: 4062-4065.

Guo L., Hu-Li J., and Paul W. E. 2005. Probabilistic regulation of IL-4 production. *J. Clin. Immunol.* 25: 573-581.

Han M. and Grunstein M. 1988. Nucleosome loss activates yeast downstream promoters in vivo. *Cell* 55: 1137-1145.

Hanlon S. E. and Lieb J. D. 2004. Progress and challenges in profiling the dynamics of chromatin and transcription factor binding with DNA microarrays. *Curr. Opin. Genet. Dev.* 14: 697-705.

Hecht A., Strahl-Bolsinger S., and Grunstein M. 1996. Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. *Nature* 383: 92-96.

Heintzman N. D., Stuart R. K., Hon G., Fu Y., Ching C. W., Hawkins R. D., Barrera L. O., Van Calcar S., Qu C., Ching K. A., Wang W., Weng Z., Green R. D., Crawford G. E., and Ren B. 2007. Distinct

and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* 39: 311-318.

Hollenhorst P. C., Shah A. A., Hopkins C., and Graves B. J. 2007. Genome-wide analyses reveal properties of redundant and specific promoter occupancy within the ETS gene family. *Genes Dev.* 21: 1882-1894.

Ioshikhes I. P., Albert I., Zanton S. J., and Pugh B. F. 2006. Nucleosome positions predicted through comparative genomics. *Nat. Genet.* 38: 1210-1215.

Jing H., Vakoc C. R., Ying L., Mandat S., Wang H., Zheng X., and Blobel G. A. 2008. Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. *Mol. Cell.* 29: 232-242.

Johnson K. D. and Bresnick E. H. 2002. Dissecting long-range transcriptional mechanisms by chromatin immunoprecipitatian. *Methods* 26: 27-36.

Jothi R., Cuddapah S., Barski A., Cui K., and Zhao K. 2008. Genomewide identification of in vivo protein-DNA binding sites from ChIP-Seq data. *Nucleic Acids Res.* 36: 5221-5231.

Kafri T., Ariel M., Brandeis M., Shemer R., Urven L., McCarrey J., Cedar H., and Razin A. 1992. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev.* 6: 705-7 14.

Keshet I., Schlesinger Y., Farkash S., Rand E., Hecht M., Segal E., Pikarski E., Young R. A., Niveleau A., Cedar H., and Simon I. 2006. Evidence for an instructive mechanism of de novo methylation in cancer cells. *Nat. Genet.* 38: 149-153.

Kim T. H. and Ren B. 2006. Genome-wide analysis of protein-DNA interactions. *Annu. Rev. Genonsics Hum. Genet.* 7: 81-102.

Klose R. J. and Bird A. P. 2006. Genomic DNA methylation: The mark and its mediators. *Trends Biochem. Sci.* 31: 89-97.

Kuo M. H. and Allis C. D. 1999. In vivo cross-linking and immunaprecipitation for studying dynamic protein: DNA associations in a chromatin environment. *Methods* 19: 425-433.

Kurdistani S. K. and Grunstein M. 2003. In vivo protein-protein and protein-DNA crosslinking for genomewide binding microarray. *Methods* 31: 90-95.

Li. Q. and Johnston S. A. 2001. Are all DNA binding and transcription regulation by an activator physiologically relevant? *Mol. Cell. Biol.* 21: 2467-2474.

Li X. Y., MacArthur S., Bourgon R., Nix D., Pollard D. A., Iyer V. N., Hechmer A., Simirenko L., Stapleton M., Luengo Hendriks C. L., et al. 2008. Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol.* 6: e27.

Lis J. and Wu C. 1993. Protein traffic on the heat shock promoter: Parking, stalling, and trucking along. *Cell* 74: 1-4.

Lomvardas S. and Thanos D. 2001. Nucleosome sliding via TBP DNA binding in vivo. *Cell* 106: 685-696.

Mardis E. R. 2007. ChIP-seq: Welcome to the new frontier. *Nat, Methods* 4: 613-614.

Mardis E. R. 2008. Next-generation DNA sequencing methods. Annu. Rev. *Genomics Hum. Genet.* 9: 387-402.

McAndrew P. C., Svaren J., Martin S. R., Horz W., and Goding C. R. 1998. Requirements for chromatin modulation and transcription activation by the Pho4 acidic activation domain. *Mol. Cell. Biol.* 18: 5818-5827.

Martone R., Euskirchen G., Bertone P., Hartman S., Royce T. E., Luscombe N. M., Rinn J. L., Nelson F. K., Miller P., Gerstein M., et al. 2003. Distribution of NF-κB-binding Sites across human chromosome 22. *Proc. Natl. Acad. Sci.* 100: 12247-12252.

Mavrich T. N., Ioshikhes I. P., Venters B. J., Jiang C., Tomsho L. P., Qi J., Schuster S. C., Albert I., and Pugh B. F. 2008. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* 18: 1073-1083.

McNally J. G., Muller W. G., Walker D., Wolford R., and Hager G. L. 2000. The glucocorticoid receptor:

Rapid exchange with regulatory sites in living cells. *Science* 287: 1262-1265.

McPherson C. E., Shim E. Y., Friedman D. S., and Zaret K. S. 1993. An active tissue-specific enhancer and bound transcription factors existing in a precisely positioned nucleosomal array. *Cell* 75: 387-398.

Meissner A., Mikkelsen T. S., Gu H., Wernig M., Hanna J., Sivachenko A., Zhang X., Bernstein B. E., Nusbaum C., Jaffe D. B., Gnirke A., Jaenisch R., and Lander E. S. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454: 766-770.

Mostoslavsky R., Singh N., Kirillov A., Pelanda R., Cedar H., Chess A., and BergmanY. 1998. Kappa chain monoallelic demethylation and the establishment of allelic exclusion. *Genes Dev.* 12: 1801-1811.

Mueller P. R. and Wold B. 1989. In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* 246: 780-786.

Mueller P. R., Salser S. J., and Wold B. 1988. Constitutive and metal-inducible protein: DNA interactions at the mouse metallothionein I promoter examined by in vivo and in vitro footprinting. *Genes Dev.* 2: 412-427.

O' Geen H., Nicolet C. M., Blahnik K., Green R., and Farnham P. J. 2006. Comparison of sample preparation methods for ChIP-chip assays. *BioTechniques* 41: 577-580.

O' Neill L. P. and Turner B. M. 2003. Immunoprecipitation of native chromatin: NChIP. *Methods* 31: 76-82.

O' Neill L. P., VerMilyea M. D., and Turner B. M. 2006. Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. *Nat. Genet.* 38: 835-841.

Orlando V. 2000. Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends Biochem. Sci.* 25: 99-104.

Parekh B. S. and Maniatis T. 1999. Virus infection leads to localized hyperacetylation of histones H3 and H4 at the IFN-β promoter. *Mol. Cell* 3: 125-129.

Patrinos G. P., de Krom M., de Boer E., Langeveld A., Imam A. M., Strouboulis J., de Laat W., and Grosveld F. G. 2004. Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. *Genes Dev.* 18: 1495-1509.

Pazin M. J., Bhargava P., Geiduschek E. P., and Kadonaga J. T. 1997. Nucleosome mobility and the maintenance of nucleosome positioning. *Science* 276: 809-812.

Phuc Le P., Friedman J. R., Schug J., Brestelli J. E., Parker J. B., Bochkis I. M., and Kaestner K. H. 2005. Glucocorticoid receptor-dependent gene regulatory networks. *PLoS Genet.* 1: e16.

Ramirez-Carrozzi V. R., Nazarian A. A., Li C. C., Gore S. L., Sridharan R., Imbalzano A. N., and Smale S. T. 2006. Selective and antagonistic functions of SWI/SNF and Mi-2beta nucleosome remodeling complexes during an inflammatory response. *Genes Dev.* 20: 282-296.

Rando O. J. and Ahmad K. 2007. Rules and regulation in the primary structure of chromatin. *Curr. Opin. Cell Biol.* 19: 250-256.

Rao S., Procko E., and Shannon M. F. 2001. Chromatin remodeling, measured by a novel real-time polymerase chain reaction assay, across the proximal promoter region of the IL-2 gene. *J. Immunol.* 167: 4494-4503.

Reddy K. L., Zullo J. M., Bertolino E., and Singh H. 2008. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* 452: 243-247.

Reik A., Schutz G., and Stewart A. F. 1991. Glucocorticoids are required for establishment and maintenance of an alteration in chromatin structure: Induction leads to a reversible disruption of nucleosomes over an enhancer. *EMBO J.* 10: 2569-2576.

Ren B. and Dynlacht B. D. 2004. Use of chromatin immunoprecipitation assays in genome-wide location analysis of mammalian transcription factors. *Methods Enzymol.* 376: 304-315.

Ren B., Robert F., Wyrick J. J., Aparicio O., Jennings E. G., Simon I., Zeitlinger J., Schreiber J., Hannett N., Kanin E., Volkert T. L., Wilson C. J., Bell S. P., and Young R. A. 2000. Genome-wide location and function of DNA binding proteins. *Science* 290: 2306-2309.

Richard-Foy H. and Hager G. L. 1987. Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. *EMBO J*. 6: 2321-2328.

Robertson G., Hirst M., Bainbridge M., Bilenky M., Zhao Y., Zeng T., Euskirchen G., Bernier B., Varhol R., Delaney A., Thiessen N., Griffith O. L., He A., Marra M., Snyder M., and Jones S. 2007. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat. Methods* 4: 651-657.

Rundlett S. E., Carmen A. A., Suka N., Turner B. M., and Grunstein M. 1998. Transcriptional repression by UME6 involves deacetylation of lysine 5 of histone H4 by RPD3. *Nature* 392: 831-835.

Sambrook J. and Russell D. 2001. *Molecular cloning: A laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Saunders A., Core L. J., and Lis J. T. 2006. Breaking barriers to transcription elongation. *Nat. Rev. Mol. Cell Biol*. 7: 557-567.

Schones D. E., Cui K., Cuddapah S., Roh T. Y., Barski A., Wang Z., Wei G., and Zhao K. 2008. Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132: 887-898.

Segal E., Fondufe-Mittendorf Y., Chen L., *Thåström* A., Field Y., Moore I. K., Wang J. P., and Widom J. 2006. A genomic code for nucleosome positioning. *Nature* 442: 772-778.

Sekinger E. A., Moqtaderi Z., and Struhl K. 2005. Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Mol. Cell* 18: 735-748.

Shemer R., Kafri T., O' Connell A., Eisenberg S., Breslow J. L., and Razin A. 1991. Methylation changes in the apolipoprotein AI gene during embryonic development of the mouse. *Proc. Natl. Acad. Sci.* 88: 11300-11304.

Simonis M., Kooren J., and de Laat W. 2007. An evaluation of 3C-based methods to capture DNA interactions. *Nat. Methods* 4: 895-901.

Simonis M., Klous P., Splinter E., Moshkin Y., Willemsen R., de Wit E., van Steensel B., and de Laat W. 2006. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* 38: 1348-1354.

Smale S. T. and Fisher A. G. 2002. Chromatin structure and gene regulation in the immune system. *Annu. Rev. Immunol.* 20: 427-462.

Solomon M. J., Larsen P. L., and Varshavsky A. 1988. Mapping protein-DNA interactions in vivo with formaldehyde: Evidence that histone H4 is retained on a highly transcribed gene. *Cell* 53: 937-947.

Southall T. D. and Brand A. H. 2007. Chromatin profiling in model organisms. *Brief Funct. Genomic Proteomic* 6: 133-140.

Spilianakis C. G. and Flavell R. A. 2004. Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat. Immunol.* 5: 1017-1027.

Spilianakis C. G., Lalioti M. D., Town T., Lee G. R., and Flavell R. A. 2005. Interchromosomal associations between alternatively expressed loci. *Nature* 435: 637-645.

Suzuki M. M. and Bird A. 2008. DNA methylation landscapes: Provocative insights from epigenomics. *Nat. Rev. Genet.* 9: 465-476.

Thåström A., Lowary P. T., Widlund H. R., Cao H., Kubista M., and Widom J. 1999. Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J. Mol. Biol.* 288: 213-229.

Thorne A. W., Myers F. A., and Hebbes T. R. 2004. Native chromatin immunoprecipitation. *Methods Mol Biol.* 287: 21-44.

Tolhuis B., Palstra R. J., Splinter E., Grosveld F., and de Laat W. 2002. Looping and interaction between hypersensitive sites in the active β-globin locus. *Mol. Cell* 10: 1453-1465.

Travers A. and Drew H. 1997. DNA-recognition and nucleosome organization. *Biopolymers* 44: 423-433.

Van Lint C., Emiliani S., Ott M., and Verdin E. 1996. Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. *EMBO J.* 15: 1112-1120.

van Steensel B. and Henikoff S. 2000. Identification of in vivo DNA targets of chromatin proteins using tethered Dam methyltransferase. *Nat. Biotechnol.* 18: 424-428.

Venkatasubrahmanyam S., Hwang W. W., Meneghini M. D., Tong A. H., and Madhani H. D. 2007. Genome-wide, as opposed to local, antisilencing is mediated redundantly by the euchromatic factors Setl and H2A. Z. *Proc. Natl. Acad. Sci.* 104: 16609-16614.

Verdin E., Paras P., and Van Lint C. 1993. Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. *EMBO J.* 12: 3249-3259.

Vermaak D. and Wolfe A. P. 1998. Chromatin and chromosomal controls in development. *Dev. Genet.* 22: 1-6.

Vogel M. J., Peric-Hupkes D., and van Steensel B. 2007. Detection of in vivo protein-DNA interactions using Dam-ID in mammalian cells. *Nat. Protoc.* 2: 1467-1478.

Walter J., Dever C. A., and Biggin M. D. 1994. Two borneo domain proteins bind with similar specificity to a wide range of DNA sites in *Drosophila* embryos. *Genes Dev.* 8: 1678-1692.

Wang Z., Zang C., Rosenfeld J. A., Schones D. E., Barski A., Cuddapah S., Cui K., Roh T. Y., Peng W., Zhang M. Q., and Zhao K. 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* 40: 897-903.

Wathelet M. G., Lin C. H., Parekh B. S., Ronco L. V., Howley P. M., and Maniatis T. 1998. Virus infection induces the assembly of coordinately activated transcription factors on the IFN-β enhancer in vivo. *Mol. Cell* 1: 507-518.

Weber M., Davies J. J., Wittig D., Oakeley E. J., Haase M., Lam W. L., and Schübeler D. 2005. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* 37: 853-862.

Weinmann A. S., Plevy S. E., and Smale S. T. 1999. Rapid and selective remodeling of a positioned nucleosome during the induction of IL-12 p40 transcription. *Immunity* 11: 665-675.

Weinmann A. S., Yan P. S., Oberley M. J., Huang T. H., and Farnham P. J. 2002. Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev.* 16: 235-244.

Weinmann A. S., Mitchell D. M., Sanjabi S., Bradley M. N., Hoffmann A., Liou H. C., and Smale S. T. 2001. Nucleosome remodeling at the IL-12 p40 promoter is a TLR-dependent, Rel-independent event. *Nat. Immunol.* 2: 51-57.

Yu L., Liu C., Vandeusen J., Becknell B., Dai Z., Wu Y. Z., Raval A., Liu T. H., Ding W., Mao C., Liu S., Smith L. T., Lee S., Rassenti L., Marcucci G., Byrd J., Caligiuri M. A., and Plass C. 2005. Global assessment of promoter methylation in a mouse model of cancer identifies ID4 as a putative tumor-suppressor gene in human leukemia. *Nat. Genet.* 37: 265-274.

Zhang X., Yazaki J., Sundaresan A., Cokus S., Chan S. W., Chen H., Henderson I. R., Shinn P., Pellegrini M., Jacobsen S. E., and Ecker J. R. 2006. Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* 126: 1189-1201.

Zhang X., Germann S., Blus B. J., Khorasanizadeh S., Gaudin V., and Jacobsen S. E. 2007. The *Arabidopsis* LHP1 protein colocalizes with histone H3 Lys27 trimethylation. *Nat. Struct. Mol. Biol.* 14: 869-871.

网络资源

http://www.abcam.com/index.html?rid=310; http://www.abcam.com/index.

html? pageconfig=popular _ protocols：Abcam 公司 ChIP 抗体及 ChIP 方案。

http://www. chem. agilent. com/scripts/LiteraturePDF. asp? iWHID = 42986&FileName=5989-4360EN _ LO. pdf：Agilent 公司 ChIP 分析 1.2 软件。

http://www. nimblegen. com/products/lit/chip _ hd2 _ datasheet _ 2008 _ 04. pdf：NimbleGen 公司 ChIP-chip 微阵列设计说明书。

http://www. affymetrix. com/index. affx；chiponchip _ appnote. pdf：Affymetrix 公司应用笔记，ChIP-on-chip 基因调控。

10　鉴定和表征转录调控因子的结构域

要点

- 鉴定调控因子的功能结构域对于了解其作用机制很重要。
- 一级结构的计算分析对鉴定转录调控蛋白的重要区域是一个有益的出发点。
- 蛋白质-蛋白质间相互作用的研究和活性分析对于正确评定蛋白质的功能至关重要。
- 很多调控因子都是多亚基的组装物。
- 结构域交换是一种确定结构域功能的有用方法。
- 当有一个因子或家族成员的结构或计算分析做出特定实验预测时，在这种特定情况下，定点突变是有用的。
- 分析调控因子的结构域时必须考虑邻近序列依赖的相互作用。

引言和概述

在哺乳动物基因调控领域中，了解 RNA 聚合酶 II（Pol II）转录的分子细节是最具挑战性的问题之一，这种挑战归因于解析多组分大分子组装物中转录因子、DNA 和核小体之间相互作用中固有的技术局限性。为了获得这些相互作用的详细图景，必须使用突变发生界定执行特定功能的因子的结构域。用于哺乳动物细胞内结构域分析的一般策略一直是通过突变发生与细胞培养物转染及 DNA-蛋白质或蛋白质-蛋白质相互作用研究的结合，来表征转录因子。在体外，缺失某些相互作用的突变体与体内外转录速率的改变有关。但是，使用经典遗传学不能进一步表征相互作用，因而无法获得确定性的结论。哺乳动物准遗传学系统［RNA 干扰（RNAi）敲低或改变特异性系统］的新进展正在为生化研究提供支持，并且为大分子组装物内复杂的蛋白质-蛋白质相互作用的研究铺平道路。此外，染色质免疫沉淀技术（ChIP）的使用使得详细的定位研究与遗传和分子分析相结合进行。

本章概述了鉴定哺乳动物调控蛋白的功能结构域的常用方法，其中强调了一些研究得很好的实例，将重点放在结构-功能分析中所采用的概念和方法学上，特别是涉及通用转录机器和像 p300 一样的共激活因子的突变发生的经典研究。在全部或部分结构已知的情况下，基于突变发生的结果，可做出非常明确的机制性推断。

在这本书的第一版中，我们强调了对激活因子（或抑制因子）结构域的分析，同时还讨论了 TATA 框-结合蛋白（TBP）的详细的突变分析，因为 TBP 代表了当时，乃至现在使用与功能研究相结合的定点突变来研究机制的经典方法，该领域变化相当迅速。基因组学和蛋白质组学的较大进展使得简单类型的结构域分析有用，但不完全。新的激活因子/抑制因子数量庞大，而且很多共激活因子/共抑制因子是多组分复合物，这些情况使得机制的详细分析成为一项艰巨的任务。在这一章中，我们考虑了这种新的状况，并讨论了可能揭示有意义的信息的方法。

实验策略：定义结构域

功能结构域的鉴定

随着高通量基因组学和蛋白质组学的出现，调控因子的分析也呈现新的特征。在理想的情况下，当研究一种新的调控因子时，研究者期望确定该因子包含什么结构域、该因子与什么相互作用、该因子是否具有催化活性，以及它的相互作用或活化的结果是什么。尽管许多序列特异性结合蛋白像激活因子和抑制因子一样是单体和二聚体，但共激活因子和共抑制因子是多个亚基组装物，这一事实使情况变得更为复杂。分析单个蛋白要比分析复合物更简单。在一开始就摆出的一个问题是是否结构-功能分析会取得关于所研究的系统的有价值的信息或揭示关于调控机制的新见解。已经进行了许多突变研究，但都没有超越分析的起始阶段。原因是随后的阶段更加难以实施，甚至在转录领域，这些步骤始终没有被很好地界定。然而，新的研究将增加我们对机制的认识，特别是当这些研究揭示出一种新的作用模式，或者是解析被调控系统中启动子功能的生物学

的协调一致的努力的一部分。

通过蛋白质的部分水解可以确认，作为独立折叠单位的蛋白质结构域的经典定义，在现代生物学中已被比较计算的方法所取代。比较计算的方法基于氨基酸组成及与结构和功能已知的蛋白质的相似性来预测功能性。因此，当检验新的或已知的蛋白质的功能结构域时，研究者往往总是以计算机分析为起点，在这种分析中，将蛋白质与其他真核生物的同源物或与包含属于同一家族结构域的蛋白质进行系统发生比较。使用最广泛的一套计算机分析工具是由美国国家生物技术信息中心（NCBI；http:/fwww. ncbi. nlm. nih. gov/）提供的一系列广泛的数据库和搜索引擎。专题 10.1 列举了基本的结构域搜索方案和 4 种蛋白质——Zif268、p300、转移酶 G9a 和 JmjD2A 的搜索结果。

尽管这种方法很有效，但某些结构域和相互作用的表面不属于可容易确认的类型，同时与功能分析相结合的突变在鉴定这类结构域是必需的。并且，保守的结构域搜索并不总是能鉴定出短小的模体，如核受体相互作用框（LXXLL 模体）或泛素化、SUMO 化及磷酸化的假定位点。在这种情况下，必须采用其他更特异的方法。例如，ExPASy（专家蛋白质分析系统，Expert Protein Analysis System）网站（http://www. expasy. org/）包含了一系列用于蛋白质结构域和功能分析的在线工具的链接。在“蛋白质组学和序列分析工具(proteomics and sequence analysis tools)”部分包含有进行保守结构域和较小模体分析的链接［即“模式和图谱检索（patterns and profile searches)”］。

结构域分析从概念的角度来看是相对简单的，即使某些方法技术要求高，并且有很多局限性。如第 11 章所讨论的，DNA-结合因子中结构域的鉴定已被大大简化。得知某一因子含有锌指结构域、同源结构域或碱性螺旋-环-螺旋（bHLH），即可将其归入一个蛋白质家族，该蛋白质家族的同源性和结构域组织是可比较的，因为结构通常是可得到的。此外，许多修饰或结合染色质的酶也被鉴定，并且根据它们的结构和功能进行了分类。然而，一个更成问题的情况是该因子是多亚基组装物的组分。在这种情况下，可以通过删除突变或去除单个的结构域来解决，但该效应的结果可能不是直接的。此外，在写这篇文章的时候，几乎没有大的多亚基调控因子的结构是已知的，RNA Pol II 可能是显著的例外（Gnatt et al. 2001）。下面我们会讨论这些问题中的一部分。

专题 10.1　结构域的计算分析

调控因子中的结构域鉴定起始于使用基于网络服务器的计算方法。我们将描述一个非常基本的搜索，同时提醒大家有许多完成此种搜索的生物信息学资源（见 http://harvester. fzk. de）。

1. 要获得蛋白质序列，登陆 www. ncbi. nlm. nih. gov。点击“GenBank”，然后在“Search”下选择“Protein database”，输入要分析的蛋白质名称。以 p300 为例，点击“Search”后将会出现来自各种各样来源和物种的 p300 的名单。

2. 使用 E1A 结合蛋白 p300［人源］，序列登录号为 NP _ 001420。点击 NP _ 001420，下载序列、参考文献列表及对序列特征的描述。

3. 在“ORIGIN”部分找到氨基酸序列并且复制序列。

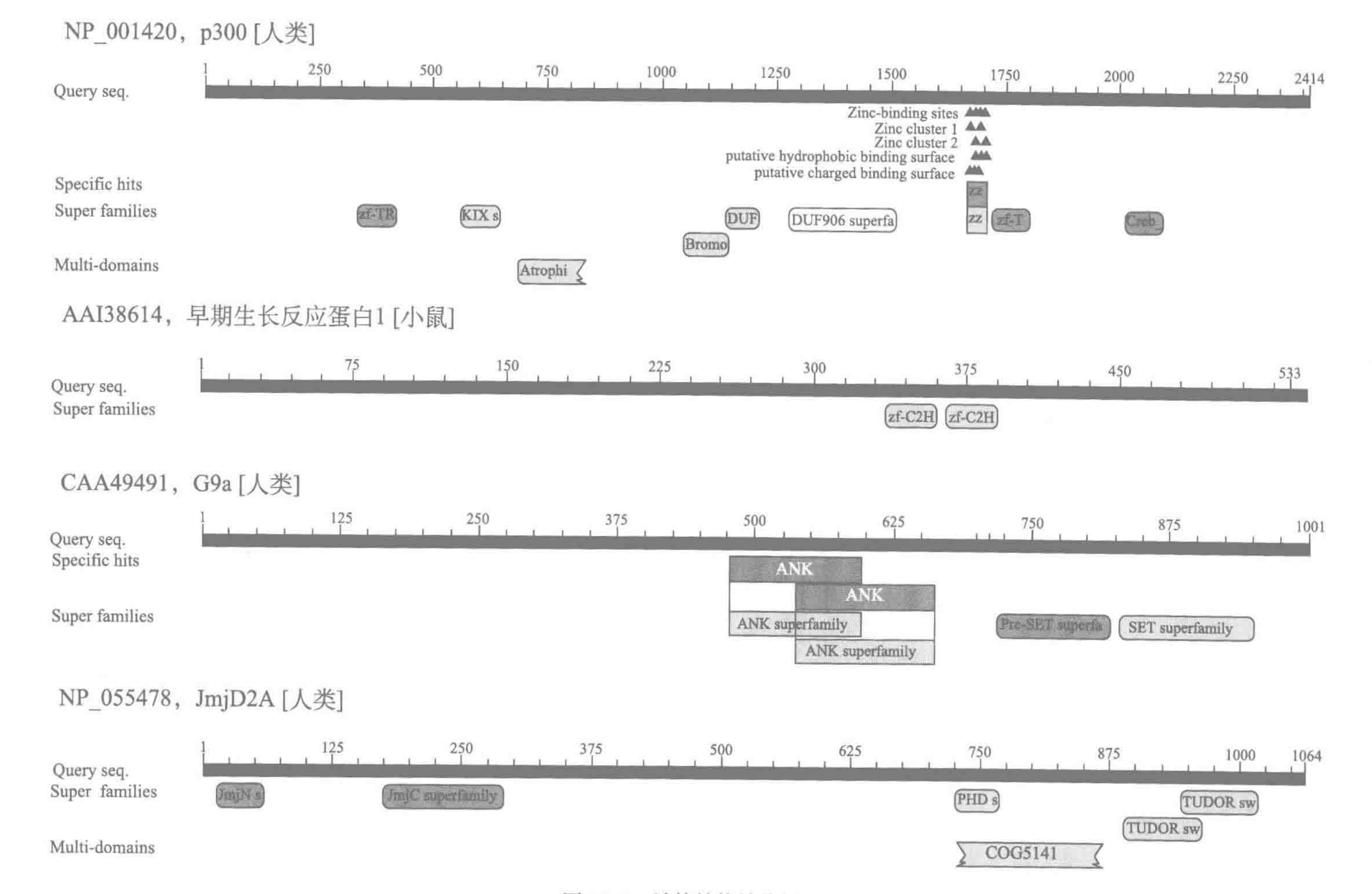


图 10.1 计算结构域分析。

4. 在这个阶段，返回到 www. ncbi. nlm. nih. gov，并且点击顶部菜单栏的“BLAST”按钮。在“Specialized BLAST”标题下点击“Find Conserved Domains in Your Sequence (cds)”。在输入框内粘贴蛋白质序列。

5. 通过下拉菜单选择要搜索的数据库，默认为“CDD”，点击“Submit”。该服务器将下载一个页面，该页面带有结构域的图形显示和结构域搜索命中的名单并附相应的链接。

作为示例，图 10.1 显示了对 p300 [组蛋白乙酰转移酶（HAT）和共激活因子]、Zif268（锌指 DNA-结合蛋白）、G9a（组蛋白甲基化酶）及 JmjD2A（组蛋白去甲基化酶）的检索结果。在浏览网页时，你会注意到结构、参考文献和相关蛋白质的链接。我们建议作为一个初学者，应在一开始就探索这些选项来理解数据库的全部功能。

基本突变原理

缺失突变是研究调控因子功能结构域的最为常用的工具之一。使用各种各样的方法从基因去除编码序列，包括 PCR 连接、限制酶切割及外切核酸酶，如 Bal31 等（详细方案见 Sambrook and Russell 2001，第 13 章）。在从比较分析中获知目的结构域的情况下，可将高选择性缺失和点突变引入功能研究中。可将缺失做成重组形式（见专题 10.2）用于生化研究，或者将缺失瞬时地或稳定地引入细胞培养物，以研究对体内功能的影响。专题 10.2 简要概述了为了各种各样的应用，如何表达调控因子。

专题 10.2　表达系统

本章探讨了结构域分析的数个不同阶段，这种结构域分析需要呈重组和体内形式的两种蛋白质的表达，以用于表达分析。在这本书的前一版中，我们用整个章节介绍了生产重组蛋白的方法。自那以来，该领域进展显著，并且用于蛋白质表达的大多数信息和系统可以普遍地通过商业资源获得。然而，这可能令那些初涉该领域的研究者难以找到出发点。在这里，我们简要地概述广泛应用于该领域的表达系统的类型和一些网络资源，在选择表达系统时作为指导。

用于瞬时分析的真核表达系统

为了用报告基因进行功能分析，对调控蛋白的结构域的分析通常使用载体，该载体中将编码调控因子的 cDNA 置于组成型启动子（即巨细胞病毒，CMV）或是诱导型启动子的控制之下。相对于像杆状病毒一样的表达系统（见下文），在这些分析中合成的蛋白质的量不高。但是，在细胞与细胞基础上，如通过 CMV-驱动的 cDNA 产生的蛋白质的量要显著高于内源性水平，在解释数据时必须考虑这一点。这些载体通常包含可选的标记，可用于分离载体已经稳定整合到基因组的细胞克隆。在表达表位标记的蛋白质用于纯化蛋白质复合物时，通常采用这种类型的方法。在这种情况下，也可以将 cDNA 引入有利于整合的反转录病毒载体中。有多家公司出售稳定的和可诱导的系统（基于质粒或反转录病毒）。也可用小规模转染制备可被标记和纯化的重组蛋白。或者，如果蛋

白质过表达，单独的提取物可用于这样的应用，如电泳迁移率变动分析（EMSA）和谷胱甘肽-*S*-转移酶（GST）融合捕获分析（关于这种载体的来源见网络资源）。注意，有数个表达水平显著高于 CMV 载体的反转录病毒和腺病毒载体，有时也使用这些载体。

杆状病毒系统

这一系统在该领域中被广泛使用，用于表达单个蛋白质和较小复合物，以便用于生化分析。基本的杆状病毒程序涉及将 cDNA 克隆进一个穿梭载体，该载体含有通常是多角体蛋白启动子的强杆状病毒启动子和侧翼序列。然后，产生的质粒在大肠杆菌（*E. coli*）或昆虫组织培养细胞（通常是 Sf9 细胞）中与病毒基因组重组。病毒扩大培养，并用于感染昆虫细胞。因为昆虫细胞已适于在旋动培养中生长，使得在理想条件下大规模感染时，每升培养基中可以产生 1～10 mg 的蛋白质。扩大培养前，必须进行时间进程分析，以确定最佳的诱导动力学条件。制备标准的核提取物和细胞质提取物，并使用 His6 标签从这些提取物中纯化蛋白质（见专题 10.3）。有很多这个系统的变种，读者应该通过浏览网络资源中列出的站点决定哪种系统最适合于自己的需要。

原核系统

该系统包括但不限于大肠杆菌（*E. coli*）中的噬菌体 T7 和 Tac 表达系统。这些系统非常适合合成大量的（即 10～100 mg）分子质量小到中等大小（<75 kDa）的、不需要翻译后修饰的重组蛋白。大肠杆菌（*E. coli*）对于表达用于特定研究的蛋白质的小结构域非常有用，如 DNA-结合结构域或 GST 融合。Tac（“Trp-Lac”）启动子是 Trp 启动子的－35 一致序列区域（含序列 TTGACA）和 Lac UV5 启动子的－10 一致序列区域（含序列 TATAAT，也称为 Pribnow box；Amann et al. 1983；de Boer et al. 1983）的融合片段。这种特殊的序列组合对 *E. coli* RNA 聚合酶具有高亲和性，并且一旦从多拷贝质粒诱导，可合成大量的细胞 mRNA。

在正常对数生长期的细胞内，作为 *lacI* 基因产物的 Lac 阻遏蛋白可抑制启动子。LacI 蛋白紧密结合到位于启动子内的它的操纵子上，但是 LacI 蛋白可被乳糖、半乳糖及在野生型 *E. coli* 中发现的其他代谢物，或者乳糖类似物异丙基-β-D 硫代半乳糖（IPTG）解离。cDNA 被克隆在质粒 Tac 启动子的下游，并且在生长到中晚对数期的转染 *E. coli* 中，通过 IPTG 可诱导 cDNA 的表达。为了保持从 Tac 载体诱导表达的能力，必须在含有 *lacIq* 等位基因的菌株中进行 cDNA 克隆；*lacIq* 等位基因是 *lacI* 基因的过表达等位基因（Brent and Ptashne 1981）。使用该等位基因过量表达 LacI 蛋白可防止 Tac 启动子的渗漏表达；Tac 启动子的渗漏表达是有害的，并且会影响质粒的稳定性和完整性。表达 GST 融合蛋白的 pGEX 载体（GE Heaithcare）含有 Tac 启动子，但 Tac 系统的许多变体可以通过学术性资源获得。

由 Studier 及其同事发展的 T7 RNA 聚合酶系统被认为在 *E. coli* 中具有比 Tac 或其他系统更高的蛋白质生产效率（Studier and Moffatt 1986；Rosenberg et al. 1987；Studier et al. 1990）。T7 系统包含两个组分。第一个组分是一个特定的 *E. coli* 菌株，该菌株带有 Lac 启动子控制下的 T7 RNA 聚合酶基因，该基因通常通过噬菌体溶源体整合到基因组中。第二个组分是含有 T7 启动子和终止子的质粒。cDNA 被克隆在 T7 启动子下游的多克隆位点。为了避免上述的 Tac 系统中渗漏表达的可能性，在 T7 聚合酶缺失

的菌株中进行克隆，然后将质粒导入适当的菌株中［如 BL21 (DE3)］。IPTG 的添加导致 T7 RNA 聚合酶的合成，T7 RNA 聚合酶依次结合表达载体上的启动子并转录 cDNA。T7 是非常强的 RNA 聚合酶，可合成大量 mRNA，结果导致非常高水平的蛋白质过表达。然而，正如在 Tac 系统中，在未诱导的细胞中，即使是非常低水平的表达，T7 RNA 聚合酶也可以合成足够的 mRNA，以致干扰细胞的生长和生存。有数个对付在 T7 表达系统中渗漏表达的策略，包括表达 T7 RNA 聚合酶抑制剂（T7 溶菌酶）或 LacIq 的特殊菌株。详细信息查阅网络资源中列出的网络出处。

体外转录/翻译反应

体外转录/翻译反应在试管中进行，并且采用补充了外源性氨基酸的翻译提取物翻译合成的 mRNA，合成的 mRNA 由源于噬菌体的 RNA 聚合酶（T7、SP6 和 T3）通过体外转录生成。体外系统可以支持大分子质量的重组因子的合成，但合成的量比较低（即 1～100 ng）。在体外系统中，cDNA 一般克隆到含有噬菌体 T7、T3 或 SP6 启动子的载体中。许多公司出售合适的载体。例如，Promega 公司出售 pGEM 系列的载体，该系列载体包含有会聚的 T7 和 SP6 启动子，但两者被多克隆位点分开。然后，用噬菌体 RNA 聚合酶生成合成的 mRNA，再将合成的 mRNA 添加到翻译提取物中。在某些商业性获得的试剂盒中，也可以将转录和翻译结合起来。这些系统通常产生 10～100 ng 的蛋白质，但这一蛋白质的量可足够用于小规模 EMSA 分析。这些系统也用于生产少量的^{35}S 标记蛋白，用于像结合到 GST 融合蛋白一样的结构域分析。通过使用带帽的 mRNA 或加入不依赖于帽子的翻译元件（CITE），如脑心肌炎病毒非翻译前导序列，可改善翻译效率。这些修改已经用于产生大量的适用于广泛的 DNA 结合和二聚体分析的蛋白质（McStay et al. 1991）。体外翻译系统的一个次要问题是，如果生产分子质量大的蛋白质时，体外翻译系统有生产截短产物的趋势，这种截短产物来自内部甲硫氨酸处的翻译起始。因此，当产生的蛋白质以放射性氨基酸，如［^{35}S］甲硫氨酸和半胱氨酸标记时，内部翻译产物往往显示为低分子质量产物的梯状条带，而加帽和 CITE 的使用可将这一问题降到最低限度。

缺失突变存在着很多局限性，包括蛋白质是多亚基复合物的组分的可能性，以及缺失会影响亚基间的相互作用或复合物的完整性的可能性。此外，缺失会经常导致蛋白质的错误折叠或不稳定。在最基础的水平上，应当做尝试，以确保缺失突变体在导入培养细胞后的确被合成。应权衡使用针对天然蛋白的抗体，防备诱变可能除去针对多克隆或单克隆抗体的主要活性表位的可能性。通常，为了促进蛋白质的检测，在蛋白质的氨基和（或）羧基末端引入各种各样的表位标签（即 FLAG、Myc 或 HA），以探测转染细胞中突变体的水平。专题 10.3 描述了蛋白质检测和（或）纯化中常用的标签。必须权衡添加表位的决策，因为有可能表位本身能影响功能。例如，如果羧基末端或氨基末端的缺失影响功能，那么，研究者应该确保处于缺失终点的表位标签不会促成活性的获得或失去。如果该表位标签不影响功能，则该附加表位可用于免疫亲和树脂，以分离蛋白质或缺失突变体，以及任何相关蛋白质，以便用于生化分析。

专题 10.3 标签识别与检测的通用方法

组氨酸标签

为促进在细菌或真核系统中过量表达的蛋白质的纯化，许多表达载体在克隆 cDNA 的多聚接头的 5′ 端 或 3′ 端编码 6 个串联的组氨酸（His6）。然后，含有融合蛋白的裂解物可以在复合了镍-次氮基三乙酸（Ni-NTA）的 Sepharose 或硅层析柱上分级分离。QIAGEN 公司出售镍-纯化柱试剂盒。表达蛋白上的组氨酸结合树脂内的金属离子，从而使得蛋白质附着到层析柱的基质上。来自 Clontech 公司的钴基-Talon 金属树脂声称可产生更好的纯化结果，因为这种树脂可只结合毗邻的组氨酸残基，而镍可结合非毗邻、附近的组氨酸残基。大多数杂蛋白不与镍-Sepharose 结合而穿过层析柱。然后，使用咪唑缓冲液（或者根据实验方案，用盐酸胍缓冲液，pH=5）将已附着的 His6 蛋白从层析柱上洗脱下来。带电荷的咪唑和组氨酸竞争性地与金属结合，并且破坏了组氨酸与金属的相互作用，因此从树脂上将蛋白质洗脱下来。His6 标签也是许多真核载体，特别是杆状病毒表达系统的组件。真核提取物中增加的蛋白质复杂性意味着某些蛋白质具有附着到树脂上的天然的组氨酸片段。在这种情况下，从镍层析柱上洗脱的蛋白质通常比较粗糙，并且必须进一步纯化（在镍-Sepharose 亲和层析之前或之后）。His6 标签的主要优点是层析可以在变性条件下进行（即 6～8 mol/L 尿素），从而促进在 *E. coli* 或其他表达系统中不溶性蛋白质的纯化。

GST 融合蛋白

谷胱甘肽-*S*-转移酶（GST）是一个 26 kDa 的蛋白质。在最常用的载体中（那些来自最初的 Amersham Pharmacia Biotech pGEX 系列的载体，现在来自 GE Healthcare），将日本血吸虫（*Schistosoma japonicum*）GST 融合到由 *lacIq* 控制的 Tac 启动子表达的 cDNA 上。IPTG 的添加抑制了 *lacIq* 抑制的启动子，并且将含有融合蛋白的被诱导细胞提取物通过谷胱甘肽-Sepharose 或琼脂糖层析柱（或者结合可以分批完成）。融合蛋白的 GST 部分附着到基质上，并且可通过加入 5 mmol/L 游离的谷胱甘肽洗脱。谷胱甘肽是酸性的，并且洗脱需要弱碱性或中性的 pH，所以必须小心以确保洗脱缓冲液已被适当地缓冲。该融合蛋白也可以通过位点特异性的蛋白酶切割下来（见下文）。

除了应用于纯化外，GST 还有其他优势。例如，GST 显然对某些蛋白质有稳定作用。我们已发现在 *E. coli* 中自行合成时并不能达到检测水平的若干蛋白质片段在与 GST 融合后可大量合成。此外，某些在溶液中不溶的蛋白质片段由于它们与 GST 的融合变得可溶。已经报道不同来源的 GST 可形成二聚体（Singh et al. 1987；Smith and Johnson 1988；Tiu et al. 1988；Lai et al. 1989），并且对重组因子二聚化的任何研究应考虑到这一现象。GE Healthcare 公司提供了一种用于 GST 纯化的完整的试剂盒和参考手册，包括广泛使用的 pGEX 系列载体，这些载体含有 Tac 启动子、*lacIq* 基因，以及 PreScission、凝血酶或 Xa 因子蛋白酶切割位点。

蛋白酶切割序列

凝血酶（Leu-Val-Pro-Arg-Gly-Ser）、肠肽酶（Asp-Asp-Asp-Asp-Lys）和 Xa 因子

(Ile-Glu-Gly-Arg) 等切割位点，往往放在两个融合搭档之间，如 His6 或 GST 标签与重组蛋白之间。烟草蚀纹病毒（TEV）蛋白酶切割位点（Asp-Arg-Leu-Tyr-Phe-Gln-Gly）也已被使用。这些位点使转录因子能轻松地从融合搭档（即 GST）中分离出来。当融合蛋白附着在层析柱基质上时，可将蛋白酶与融合蛋白孵育，洗脱后的蛋白质最终处于上清液中，并且通过离心或过滤直接从层析柱基质中分离出上清液，从而分离蛋白质。

免疫标签

免疫标签已成为一种在来自表达细胞的粗提取物中纯化和鉴定融合蛋白的便利且高度特异性的方法。免疫标签对于从真核表达系统中纯化蛋白质尤其有用，因为在真核表达系统中蛋白质的复杂性远远高于 *E. coli* 中。由于单克隆抗体对其表位的高亲和性和特异性，洗涤步骤可以在严紧的条件下进行，在这样的条件下免疫亲和树脂上几乎不结合杂蛋白。免疫耗竭所必需的树脂量必须依靠经验确定，因为粗提物对标记蛋白的结合有温和的抑制作用。免疫标签很少用于 *E. coli* 表达系统，因为抗体树脂价格昂贵，并且在像 GST 或 His6 等的其他标签可广泛获得时，抗体树脂在纯化上几乎提供不了什么优势。然而，在缺乏天然蛋白质抗体的情况下，免疫标签对纯化期间或实验程序中的检测会非常有用。

流感病毒血凝素（HA）标签（Tyr-Pro-Tyr-Asp-Val-Pro-Asp-Tyr-Ala）和 FLAG 标签（Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys）是融合到重组蛋白氨基末端或羧基末端的短肽表位［其他还包括水泡性口炎病毒（VSV）和 Myc 标签］。这些标签特异并紧密地结合到单克隆抗体亲和树脂上，并且可用相应的多肽高效地洗脱下来。编码 HA、VSV 和 Myc 标签的哺乳动物表达载体、这些标签相应的免疫亲和树脂，以及用于洗脱的多肽都可通过 Roche Molecular Biochemicals 公司获得。FLAG 标签树脂和多肽可以通过 Sigma-Aldrich 公司获得。一般来说，12CA5（HA）或 M1/M2（FLAG）抗体与蛋白 A-Sepharose（或琼脂糖）或者活化的亲和树脂，如溴化氰（CNBr）活化的Sepharose 微珠共价交联。然后，含有融合蛋白的细胞提取物通过柱层析或批层析穿过亲和树脂，该树脂用高盐缓冲液大量地洗涤以除去非特异性结合蛋白。结合蛋白以约 1 mg/ml 多肽及中到高盐浓度洗脱，有时，包带有多拷贝 Tag 标签的多肽可用于洗脱。FLAG Ml 抗体的结合需要 Ca^{2+}，并且通过 EGTA 对它的去除造成抗体从 FLAG 表位解离，从而为洗脱提供了一种更便利的方法。

串联亲和纯化标签

这种策略的原型产生于蛋白质组学研究，在这种研究中，研究者曾寻找一种方法，以快速纯化蛋白质复合物，从而用于大规模蛋白质组学分析。最初的串联亲和纯化（TAP）标签包括钙调蛋白结合肽、TEV 切割位点，以及两个金黄色葡萄球菌（*Staphylococcus aureus*）蛋白 A 的免疫球蛋白 G（IgG）结合结构域（Rigaut et al. 1999；Puig et al. 2001）。TAP 标签通常附加在蛋白质氨基末端或羧基末端。在第一步中，含有该蛋白质的提取物首先分批与 IgG-Sepharose（GE Healthcare 17-0969-01）孵育，再洗涤去除未结合的蛋白质，使用 TEV 蛋白酶（Invitrogen 12575-015）切割法洗脱结合的蛋白质，TEV 蛋白酶的酶切位点是 ENLYFQG。在第二步中，该蛋白质在钙离子的存在下与钙调蛋白-Sepharose

(GE Healthcare 17-0529-01) 结合、洗涤，并且之后用含有钙螯合剂 EGTA 的缓冲液洗脱。该方法的缺点是切割后钙调蛋白结合肽和 TEV 切割位点仍然保持连接，结果在蛋白质末端留下约 45 个氨基酸。此外，蛋白酶切割并不总是完全的。已经构建了许多建立在先前存在的标签上的附加 TAP 标签，包括 FLAG-HA (Sigma TP0010) 和 His6-FLAG。这些标签的优点是它们明显更小，并如上所述不太可能妨碍其附着的蛋白质的特性。此外，这些标签使用标准的洗脱技术和树脂。

生物素标签

生物素标记策略一般用于转录复合物的蛋白质组学分析。将编码 15～23 个氨基酸标签 (de Boer et al. 2003; Rodriguez et al. 2006) 的一段 DNA 片段添加到 cDNA 上并构建稳定表达的细胞系。其中，该氨基酸标签包含生物素化靶序列 MASSL-RQILDSQKMEWRSNAGGS (生物素发生在 K 上)。然后，将编码 *E. coli* BirA 蛋白的 cDNA 表达载体转染该细胞系或构建稳定的细胞系 (见 Wang et al. 2006a)。BirA 蛋白生物素化上面指出的赖氨酸。然后，带有生物素化蛋白的细胞提取物结合到链亲和素偶联的微珠上，这一相互作用的高亲和力 (K_d 约为 10^{-15} mol/L) 允许对微珠进行相对严紧的洗涤，以去除未结合的蛋白质。然后，通过在含有十二烷基硫酸钠 (SDS) 的缓冲液中煮沸，洗脱微珠，并且在 SDS 凝胶中分级分离洗脱的蛋白质。对于蛋白质组学分析来说，该方法的主要优点是所制备的蛋白质是高纯度的，但缺点是在自然条件下难以去除蛋白质，除非在构体中加入蛋白酶切割序列；在这种情况下，可以通过蛋白酶处理洗脱蛋白质 (见上文)。通过添加另一个标签，如 FLAG，并且在链亲和素结合步骤前用 FLAG 肽洗脱，这种方法也可适用于串联亲和纯化。

心肌激酶标签

心肌激酶 (HMK) 标签不会对纯化带来好处，尽管它们对重组蛋白的检测和操作有用。HMK 是蛋白激酶 A 的催化亚基，该酶识别和磷酸化一个位于高度特异性的五氨基酸序列 (Arg-Arg-Ala-Ser-Val) 中的丝氨酸。可将该序列融合到编码该 cDNA 的基因的 5′ 端或 3′ 端，这样产生的蛋白质将含有该序列。这个标签使研究者能够使用 HMK (Sigma, P2645) 和 [γ-^{32}P] ATP 标记纯蛋白质的羧基末端或氨基末端。标记的蛋白质有很多应用，包括作为探针筛选表达文库。

定点突变是仅改变蛋白质内确定的单个或成簇的氨基酸的方法。定点突变分析用于鉴定结构域中对功能必需的精确的氨基酸。定点突变还可用于进一步刻画通过缺失鉴定的区域，或者与晶体结构，或者与序列比较信息相结合，测试功能的模型。

一般来说，蛋白质内定点突变有两个广泛使用的变体：丙氨酸置换和基团置换。丙氨酸是一种带有甲基侧链的小疏水性氨基酸，在丙氨酸置换中，丙氨酸的密码子被引入替换原有的氨基酸 (图 10.2，比照图 10.2A 和图 10.2B) (Cunningham and Wells 1989)。丙氨酸缺少除了侧链 C_β 外的碳原子，并且被认为代表侧链截断的等价物。当置于蛋白质的溶剂暴露面上时，丙氨酸被假定对结构几乎没有整体上 (即热力学中性) 的影响。某些情况下，接触丧失突变的热力学后果可能太小，不足以产生表型。在这些情况下，就要使用基团置换。在基团置换中，插入带有突出侧链的化学上不同的氨基酸替

代野生型残基（如精氨酸变成天冬氨酸）（图 10.2，比照图 10.2A 和图 10.2C）。该概念是插入一种氨基酸，该氨基酸既能消除最初的接触，又能在空间上或化学上破坏（即电荷排斥相互作用）任何其他蛋白质与突变表面的结合。除丙氨酸和基团置换外，某些定点突变方法已被用于研究磷酸化。例如，用谷氨酸替代丝氨酸通常被认为是模拟磷酸化的丝氨酸（图 10.2D）。此外，用谷氨酰胺替代赖氨酸通常被认为是模拟乙酰化赖氨酸。虽然有些功能研究支持这些替代的有效性，但结构分析并非总是支持这些替代。更复杂的程序，如在特定位置放置半胱氨酸并对其进行化学修饰可以用于引入特殊类型的修饰（见第 13 章）。重要的是，Schultz 及其同事已经生产了一种带有独特密码子和相应的 tRNA/氨酰-tRNA 合成酶对的 *E. coli*，使用琥珀无义密码子，这种 *E. coli* 允许超过 30 种不同的非天然氨基酸掺入蛋白质中。该系统可用于生成具有特定人造氨基酸的蛋白质，这种蛋白质可用于蛋白质的结构-功能分析（Wang et al. 2006a）。

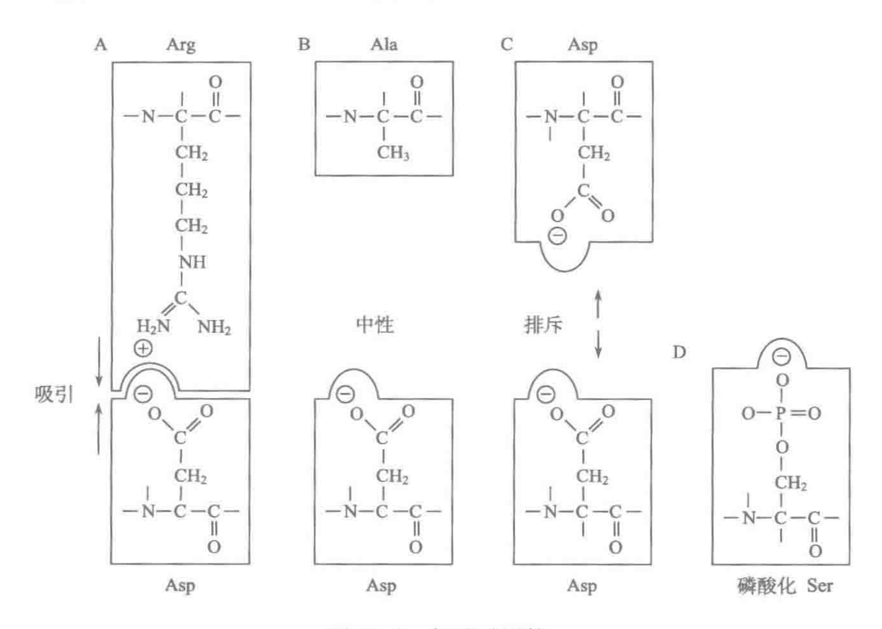


图 10.2　氨基酸置换。

当与晶体结构相结合或用于诱变相关因子的保守结构域时，定点突变是最强大的。例如，如果一个 DNA 结合结构域（见第 11 章）的晶体结构是已知的，那么可以使用定点突变来改变关键性氨基酸以证实其在溶液分析或在体内的重要性。或者，如果一个相关转录因子的结构是已知的，该因子的晶体结构就可用于指导相关蛋白质的突变研究。通过诱变在相关蛋白质中占据相似位点的氨基酸残基，研究者可以确定一些结合模式是否相似。这种情况下，往往有必要利用晶体结构中的 α-碳骨架走向建立该蛋白质的模型。结构计算程序，如 Sternberg 实验室（http://www.sbg.bio.ic.ac.uk/~phyre/）发展的 Phyre（Protein Homology/Analogy Recognition Engine，蛋白质同源/类似识别引擎），允许建立这样的模型。Phyre 程序基于相关蛋白质的已知结构进行基本结构的预测。该程序进行能量最小化，并且生成模型，生成的模型可避免空间冲突，

这些空间冲突归因于已知结构的蛋白质与同源蛋白质之间氨基酸的差异。更复杂的分析，包括蛋白质数据库文件操作（Protein Data Base file [. pdb]），可在 Coot 程序中进行（Crystallographic Object-Oriented Toolkit：http://www.ysbl.york.ac.uk/～emsley/coot/）。在缺乏结构的情况下，蛋白质的序列保守性常常可用于鉴定重要的氨基酸。例如，两个相关的调控蛋白可能共享执行关键功能的序列。如果已知一个蛋白质中的突变体能失活结构域，可以使用相似的突变体来研究另一个家族成员的功能。由于这种方法在研究结构-功能关系方面发挥的重要作用，我们已在本章最后列出了一种简单且直接的定点突变的方案。

序列特异性调控因子的结构域

可以将典型的 DNA-结合激活因子或抑制因子分成数个功能结构域，每个结构域还可进一步细分。通常，这些结构域中的一种将蛋白质靶向至合适启动子环境中的 DNA 位点，而另一种则通过与共激活因子/共抑制因子及染色质重塑/修饰因子的相互作用介导转录的激活或抑制。然而，其他结构域依赖于细胞环境发挥作用。功能结构域可以是相同多肽的一部分，或者它们可能存在于一个复合物的不同亚基上。在某些情况下，两种功能紧密地混在一起，并且只能通过详细的定点突变来刻画。MyoD 和糖皮质激素受体中的某些激活功能就是这样，结合位点的 DNA 序列可改变这些激活功能（Schena et al. 1989；Davis et al. 1990；Lefstin et al. 1994；Lefstin and Yamamoto 1998；van Tilborg et al. 2000），这一问题将在下面作简短的讨论。

DNA 结合结构域通常可进一步细分为两个或多个不同的亚结构域（图 10.3）。一个亚结构域负责识别特定的 DNA 位点并将蛋白质拴到 DNA 位点上，另一个亚结构域调节寡聚化作用（即二聚体、四聚物等），其他的亚结构域可介导与附近激活因子的协同结合。正如第 1 章和第 11 章所讨论的，正是这些相互作用的总和使得激活因子结合到适当启动子环境中的正确序列上。因为许多这些 DNA-结合结构域都属于不同的家族，这些家族的结构已通过 X 射线结晶学或核磁共振（NMR）被解析，所以我们对这些 DNA-结合结构域的结构有较多的了解（见第 1 章和第 11 章）。

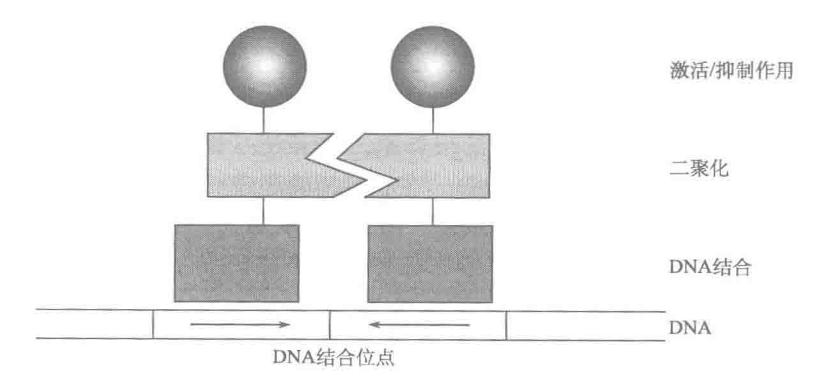


图 10.3 典型激活因子的结构域。

此外，激活/抑制结构域通过与共激活因子和共抑制因子，或者与其他修饰或重塑染色质的辅因子的直接相互作用负责影响转录。通常来说，共激活因子和共抑制因子含有修饰染色质（即 p300 或 MLL/SET1 复合物）或通用机器（即 CDK8）的催化活性。虽然序列特异性激活因子带有一个或多个单独的激活结构域，但它可能与增强体内的其他激活因子相联合协同发挥作用（见第 1 章）。与增强体装配（如协同性 DNA 结合）有关的结构域在某些突变分析（如当用野生型启动子作为报告模板时）中，可能会被错误地解释或被打分为激活结构域，因为它们的去除可能会废除增强体的组装，并且因此影响转录。设计详细的生化实验和转染分析以比较天然的和模式启动子上的突变体，对辨别这些可能性是必要的。

不同的激活因子在溶液中稳定地相互作用也是有可能的。例如，Orkin 及其同事发展了一种方法，在该方法中，将 Nanog cDNA 以 FLAG 肽和一段作为 *E. coli* 生物素连接酶（BirA）体内生物素化底物的序列标记（见专题 10.3）。该 cDNA 在设计表达 BirA 蛋白的胚胎干（ES）细胞中低水平表达。高亲和力 FLAG 和生物素标签的使用，使得研究人员能够通过一步程序或串联程序纯化 Nanog 及其相关蛋白质。结果发现 Nanog 与多种干细胞特异性的和序列特异性的 DNA 结合蛋白相互作用，包括 Oct-4。相应地，Oct-4 和其他相关因子的相互标记确证了这些因子的确共纯化。凝胶过滤确证 Nanog 以一系列分子质量高达 1 MDa 的复合物进行迁移，远远超过了 Nanog 的 34 kDa 的分子质量（Wang et al. 2006b）。

分离序列特异性调控因子的 DNA 结合和激活/抑制结构域

一般考虑

涉及激活和抑制结构域鉴定的方法在该领域中已经被描述了近 20 年。因此，策略和技术是众所周知的，尽管现在不经常使用。的确，以天然激活结构域分析所得出的信息为基础，已构建了各种各样的人工转录激活因子和抑制因子（综述见 Mapp and Ansari 2007）。用于剖析哺乳动物系统中激活因子/抑制因子的初始策略是为正在分析的蛋白质发展反式激活/抑制及 DNA 结合分析方法。反式激活/抑制分析方法是一种相对人工的方法，但它可以在结构域分析的初始阶段中揭示重要的信息，这些信息在随后更严格的生化和生物学研究中可以被验证。在反式激活/抑制分析中，使用已知对该激活因子/抑制因子作出应答的天然启动子，或与最小启动子毗邻的含有该蛋白质多个结合位点的人工启动子，驱动一个易于测量的报告基因的表达（见第 4 章的实例，如萤光素酶、绿色荧光蛋白和氯霉素乙酰转移酶）。在必须精确测定 mRNA 合成的情况下，则可以使用引物延伸、RNase 保护、S1 核酸酶、qPCR 或其他合适的方法（见第 3 章）。然后，将野生型激活因子/抑制因子及其突变体置于一个强启动子/增强子（即 CMV）的控制下，并且与报告模板共转染到合适的细胞系。

因为突变发生可以显著改变蛋白质的稳定性，因此关键的是建立一种能测定蛋白质水平的分析方法（即免疫印迹法、免疫沉淀法或 EMSA）。在观察到一种效应的情况下，研究者必须能够确证野生型蛋白质和任何突变体是否在类似的水平上被合成。如果无法获得针对所研究的蛋白质的抗体，那么，以已知表位（即 FLAG 表位）设置免疫

标签可以提供一种通过免疫沉淀、免疫印迹或两者来检测蛋白质水平的方法（见专题10.3）。应该注意的是，转染实验并不总是与结构域分析结果相符合，因为某些因子在自然染色质环境中起作用，而其他因子发挥建筑作用，这些建筑作用在简单的报告基因分析中并不总是可检测的。

重要的是要指出，在反式激活/抑制或 DNA-结合分析中，与蛋白质突变相关的“功能缺失”并不总是暗示发生突变的区域介导该效应。理由是虽然突变能通过去除关键性氨基酸直接影响 DNA 结合或转录，但也可能引起结构上的扰动，这种结构扰动会间接性地改变活性。为了确证该突变直接影响了一种重要功能，必须进行附加实验。例如，可用结构域交换实验显示结构域独自具有 DNA 结合或反式激活/抑制能力。或者，随着对结构域结构认识的增长，将突变发生和生化实验相结合（见第 12 章）的研究会揭示关于功能的新见解。

DNA 结合

对 DNA 结合结构域的最初认识是：DNA 结合结构域是蛋白质一个特定区域，当该区域被缺失时，它会废除生化分析中蛋白质对 DNA 的结合。这种结合分析通常是 DNase Ⅰ 足迹或 EMSA（见第 11 章和第 12 章）；并且体外翻译的蛋白质、*E. coli* 或杆状病毒系统中合成的蛋白质，或转染细胞提取物中的蛋白质（见专题 10.2）都可以用于进行这种结合分析。将在体外结合 DNA 的最小多肽片段定义为最小的 DNA 结合结构（见第 11 章）。

在证实了该结构域在体外的特异性结合能力后，关键问题是这种生化情况是否精确反映细胞中发生的情况。例如，如果序列特异性的决定性因素仍然存在，二聚化结构域的缺失仍会允许 DNA 的体外结合。在某些情况下，降低结合的严紧性（即盐浓度）或提高蛋白质浓度可以补偿二聚化结构域的缺失，并且允许体外结合，但相同的蛋白质在体内却没有结合能力。

有数种用于证实结构域能够在体内结合 DNA 的方法，包括硫酸二甲酯（DMS）足迹法、染色质免疫沉淀法及 DamID 法（van Steensel and Henikoff 2000；Greil et al. 2006）（见第 9 章）。比较简单的方法是显示单独的 DNA 结合结构域在共转染反式激活/抑制分析中起完整激活因子/抑制因子的“反式显性失活”，或者抑制剂的作用。高浓度的 DNA 结合结构域和完整激活因子/抑制因子竞争启动子位点，并且降低完整激活因子/抑制因子的激活（或抑制）能力。此外，一种简洁的方法是使用结构域交换（见 Brent and Ptashne 1985）。

在一个典型的 DNA 结合结构域交换中（图 10.4），将最小的 DNA 结合结构域融合到像单纯疱疹病毒的 VP16 反式激活因子那样的异源激活结构域上。例如，可以将 VP16 的激活结构域融合到无转录活性的 GAL4 DNA 结合结构域（第 1～94 位或 1～147 位氨基酸）上，以产生 GAL4-VP16。由于异源激活结构域不太可能提供对 DNA 结合重要的蛋白质组分，由含有待测蛋白质结合位点的报告基因嵌合体引起的反式激活/抑制可能暗示该 DNA 结合结构域（在这种情况下是 GAL4）在体内能够独立识别其结合位点。应当指出，这种分析方法在该领域中已广泛使用。PubMed 文献数据库中有超过 3000 篇参考文献引用了 GAL4 和 GAL4 融合蛋白。

在某些情况下，DNA 结合结构域识别其结合位点的能力受到蛋白质的其他区域的控制，或者甚至受到其他蛋白质一起的控制。在数个研究得很透彻的例子中，另一个结构域直接或间接地掩盖 DNA 结合结构域结合 DNA 的能力，除非存在适当的配体或修饰，或除非该结构域与适当的位点或正确的蛋白质搭档发生相互作用。抑制性结构域的去除可以提高体外 DNA 结合亲和性或体内激活的潜力。关键点是这种结构域在结构域交换实验中表现不明显，并且必须使用其他方法揭示其功能。以前已综述了这些自我抑制性相互作用的例子及研究它们的方法（Pufall and Graves 2002)。

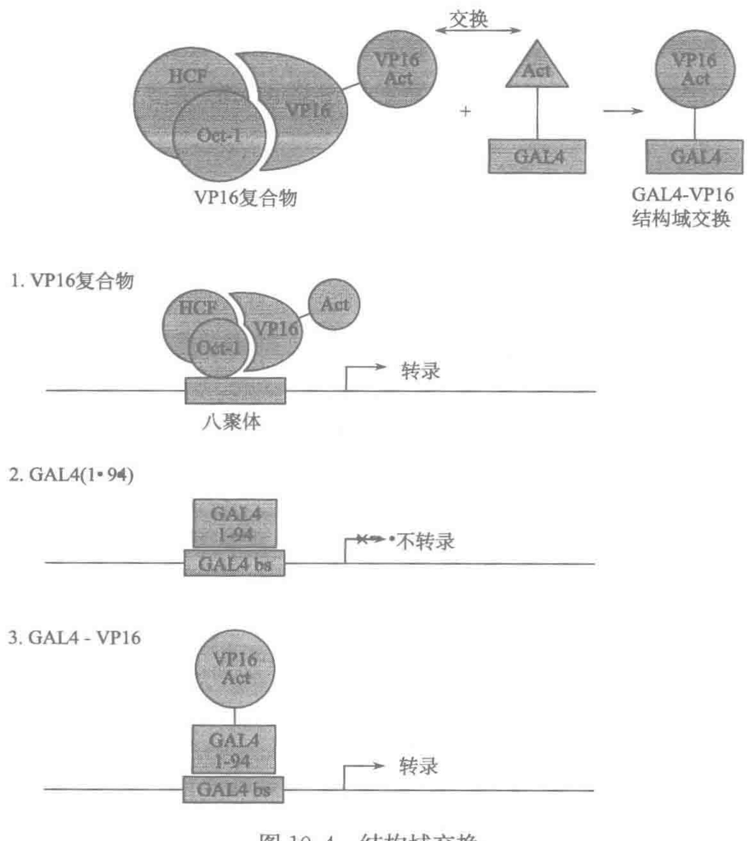


图 10.4　结构域交换。

激活与抑制

激活或抑制结构域的分析更成问题，其中必须考虑两个问题。第一，激活因子和抑制因子的普遍特征是它们可能都含有数个对调控有贡献的区域。天然蛋白质中任意一个这些区域的缺失在报告基因分析中都不足以降低激活或抑制作用。相反，转录分析中突变体的激活或抑制潜能的降低可能归因于与激活或抑制的机制无关的原因，包括蛋白质稳定性的降低或异常折叠。因此，与 DNA 结合结构域的情况不同，很多情况下对激活结构域的定义依赖于在结构域交换实验的环境中该结构域发挥功能的能力。这种通用性的方法对抑制结构域也是行之有效的，只是在这种情况下，报告基因活性的基础水平必须足够高，才能在表达中观测到可测量的降低。

在激活/抑制结构域是模块化的并且与蛋白质的其他结构域可分离的情况下，才能进行结构域交换。结构域交换要涉及构建嵌合蛋白，该蛋白质含有无转录活性的 DNA 结合结构域（即 GAL4 的 DNA 结合结构域），并且含有正在被研究的蛋白质的假定激活/抑制结构域（图 10.4）；嵌合体和报告模板共转染哺乳动物组织培养细胞，其中，报告模板含有位于最小启动子和报告基因上游的串联的寡聚化 GAL4 位点，然后将单独含有 DNA 结合结构域的调控活性与嵌合体做比较。一旦鉴定了激活结构域，就可以通过定点突变分析鉴定重要的氨基酸残基（见专题 10.4～专题 10.6 中的例子）。

专题 10.4　VP16 激活结构域：个案研究

在酵母中对激活因子的早期研究显示，激活结构域在很大程度上耐受点突变分析（Hope and Struhl 1986）。与 DNA 结合结构域中的点突变不同的是，DNA 结合结构域中的某些点突变通常具有显著的效应，而激活结构域中的点突变只表现出微弱的效应，并且只有当激活结构域的小片段被孤立研究时，才能检测到这些微弱的效应（见 Cress and Triezenberg 1991 或 Gill and Ptashne 1987）。尽管如此，回顾一下一种原型激活结构域，即 VP16 是怎样被鉴定和表征的，仍然是有益的（Triezenberg 1995，Nedialkov et al. 2003）。

VP16 是一种反式激活因子，负责激活Ⅰ型单纯疱疹病毒（HSV-1）的立即早期基因（Flint and Shenk 1997；Herr 1998）。该蛋白质是病毒衣壳的结构性组分，并且在初始感染期间，它的作用不需要病毒蛋白质的合成。在感染和病毒脱衣壳后，VP16 与细胞蛋白 Oct-1 及 HCF 形成复合物，再结合早期病毒启动子。这种结合的特异性很大程度上来自 Oct-1 与八聚体结合位点的相互作用，八聚体结合位点也称为 TATGARAT（R 代表嘌呤）；Herr 和 Sharp 实验室所做的研究已表明 VP16 的确能以序列特异的方式与 DNA 相互作用，尽管这种相互作用甚为微弱（如见 Kristie and Sharp 1990）。但是这种微小的特异性为 Oct-1 增加了足够的结合能，以致将该复合物引向病毒中的特异性八聚体结合位点，而不是细胞启动子位点。已知 Oct-1 与其他共激活因子相结合，如 B 细胞特异性蛋白 OcaB/Bob1，该蛋白质也微弱地结合 DNA，但却将 Oct-1/OcaB 复合物引向细胞启动子（Luo and Roeder 1995；Gstaiger et al. 1996）。共激活因子规划多蛋白复合体的最终特异性，这一概念在基因表达领域中是一个新兴的研究主题。和增强体的研究主题一样，这个主题为如何使用常见系列的激活因子与不同组合的搭档来影响基因表达的新模式提供了另一种解释。

由 Triezenberg、McKnight 及其同事进行的缺失分析，将 VP16 的反式激活结构域定位到由 490 个氨基酸构成的蛋白质的羧基末端 78 个氨基酸上（Triezenberg et al. 1988）。随后将该区域融合到 GAL4 DNA 结合结构域上，产生名为 GAL4-VP16 的嵌合体。从 GAL4 位点应答性的报告模板中发现，GAL4-VP16 将转录激活到高水平（Sadows-ki et al. 1988）。这证实了一种想法，即对于激活来说，含有羧基末端结构域就足够了。

后来的研究表明在嵌合体环境中的 VP16 可被进一步划分为两个亚结构域，即氨基酸 413～456 的 VP16（N）和氨基酸 457～490 的 VP16（C）。这两个结构域在病毒环境

中都具有功能，并且能够支持病毒复制，尽管比完整结构域的作用弱。最初认为这两个结构域是同一模体的简单重复片段（Emami and Carey 1992，Walker et al. 1993，Tanaka et al. 1994）。但是，由 Triezenberg 和其同事进行的详细的点突变揭示，两个亚结构域中相似位置上的氨基酸对突变的应答并不相同（Cress and Triezenberg 1991，Regier et al. 1993）。这一研究和其他的突变研究一起导致这样的提议，即两个亚结构域通过不同的机制起作用。的确，在亲和层析实验中，氨基结构域和羧基结构域表现出与不同靶标的优先相互作用（如见 Goodrich et al. 1993）。虽然该备用靶标假说有点吸引力，但是该结合研究的生理含义依然不清楚，因为氨基末端或羧基末端结构域，或其片段的重复会产生非常有力的激活因子，这些激活因子可独立于其他结构域起作用。因此，即使该结构域和不同的靶点相接触，其中这些相互作用中的任何一个对激活都是足够的（Emami and Carey 1992）。

虽然氨基和羧基亚结构域在人工启动子环境中以嵌合 GAL4-VP16 蛋白的形式是能起作用的，但它们可能起独特的启动子特异性作用，这些作用只有通过在多种启动子和细胞环境中分析该嵌合体才能体现出来。的确，在天然 HSV 启动子中，VP16 仅能从近端启动子位置激活转录。像人工系统中获得的那样（Sadowski et al. 1988，Carey et al，1990），获得远距离激活的尝试未能成功（Hagmann et al. 1997），说明该模型系统并不能精确地复述激活结构域功能的调控。

光谱学的研究已经开始揭示在其靶位点存在时 VP16 激活结构域结构转变的某些见解。使用荧光各向异性的一项研究断定，当与 TBP 相互作用时，VP16 经历了显著的结构变化（Shen et al. 1996a，b）。在这样的研究中，点突变（Cress and Triezenberg 1991）是确定结构变化特异性的必要对照。另一项研究使用 NMR 考察了 VP16 和潜在靶标 TAFⅡ32（TAF9）形成复合物时的相互作用，该研究为构象的变化是由无规卷曲转变为 α 螺旋的想法提供了依据（Uesugi et al，1997）。

Giniger 和 Rashne（1987）最初使用短肽激活结构域对激活结构域可能是两亲性螺旋的想法进行了探讨。的确，最早的激活结构域晶体结构之一，即核受体 τ 结构域的晶体结构（综述见 Mangelsdorf and Evans 1995）揭示该激活结构域为两亲性螺旋，并形成了配体结合结构域的一部分。

专题 10.5　GAL4：个案研究

为了说明 DNA 结合及二聚化问题，研究者可转向 GAL4，该蛋白质充当本章中所描述的许多策略的原型。如第 11 章中所述，在半乳糖存在时 GAL4 与酵母中半乳糖代谢基因的上游相结合，并激活这些基因（综述见 Johnston 1987；Traven et al. 2006）。最初，根本不清楚 GAL4 是否与 DNA 结合，因为 GAL4 是在一个广泛的遗传学筛选中被分离的，该遗传学筛选的目的是找到那些对半乳糖代谢重要的基因。因此，20 世纪 80 年代初，GAL4 与 DNA 结合的唯一暗示是它含有一段 DNA 序列，该序列编码一个由 Klug 及其同事发现的锌指结构的亲属蛋白（Miller et al. 1985）。由 Ptashne 及其同事用酵母野生型和突

变型 *gal4* 进行的体内足迹分析表明，GAL4 负责在 4 个 DNA 位点上产生 DMS 足迹，这 4 个位点构成半乳糖上游激活序列（UAS_G）(Giniger et al. 1985)。

这一观察促使 Ptashne 及其同事分析各种不同的 GAL4 的缺失，这些缺失是通过使用编码区内的合适限制性内切核酸酶切割位点产生的。在 *E. coli* 中作为 β-半乳糖苷酶的融合蛋白合成这些缺失体（Keegan et al. 1986)。然后，使用 DNase I 足迹法和 EMSA 分析所产生的蛋白质。通过足迹法分析，发现带有氨基末端 74 个或 147 个氨基酸的 GAL4 片段在体外结合 UAS。在一个适当报告基因上的体内分析表明，含有氨基末端但缺少羧基末端的衍生物不能激活。推断这些衍生物在体内的确能与 DNA 结合，但只是缺乏激活结构域。一系列异源激活结构域包括 VP16 与体外鉴定的这一 DNA 结合结构域的融合证实了这一推断，因为实验表明这些融合蛋白恢复了该 DNA 结合区域在体内的转录激活能力（Ma and Ptashne 1987a，b；Sadowski et al. 1988)。相反，当将羧基末端上围绕着删除区域的氨基酸融合到无转录活性的 LexA DNA 结合结构域上时，前者赋予后者在体内从 LexA 位点激活转录的能力（Brent and Rashne 1985)。这些结构域交换是真核生物激活因子模块性的最初实例，并且这些结构域交换与在 GCN4 上同时进行的实验一起，为分析真核激活因子提供了范例（Hope and Struhl 1986)。

当分析二聚化界面时，情况更为复杂。有三项实验表明 GAL4 以二聚体形式结合 DNA：17 bp 位点的二分体特性、通过化学和核酸酶足迹技术揭示的 GAL4 与该位点的对称性接触，以及最后的异源二聚体分析（见第 11 章；Carey et al. 1989)。最早的 GAL4 分析使用通过编码片段的限制酶缺失产生的突变体。DNA 结合结构域含有介于氨酸基 11 和氨酸基 38 之间的 6 个 Cys，当时这些 Cys 看起来像锌指 DNA 结合结构域。然而，通过从氨基末端缺失来进一步界定该结构域的任何尝试都废除了该蛋白质的 DNA 结合能力。因此，为了鉴定最小的 DNA 结合结构域，构建了从氨基酸 147 向 GAL4 氨基末端的进行性羧基末端缺失。早先已经注意到 GAL4 的氨基酸 1～74 对 DNA 的结合不及氨基酸 1～147 对 DNA 的结合那样紧密，因此推断位置 74 处于假设的二聚体区域中。为了弥补缺失过程中可能损失的二聚化界面，将产生的突变体融合到 λ 抑制因子的二聚体结构域上。使用这种方法，发现特异性有效结合 DNA 的最小片段是氨基酸 1～65。将单独的氨基酸 1～65 与 λ 抑制因子融合的氨基酸 1～65 的亲和力进行比较，结合显示亲和力的差异达 100 倍或更大。的确，氨基酸 1～147 的自身缺失表明，氨基酸 94 之前的缺失几乎无影响，此后，亲和力逐渐下降，直到在氨基酸 1-65 时亲和力降至最低。注意，虽然通过上面的讨论，氨基酸 65 和氨基酸 94 之间的区域可能涉及二聚化，但从未通过充分的测试对其进行确证［即证明在嵌合蛋白的环境中具有功能，并且赋予单体形成二聚体的能力（即类似于 Hu et al. 1990 的方法)］。然而，最终，GAL4 的结构揭示该二聚化结构域由 3 个 α 螺旋构成。第一个 α 螺旋出现在 GAL4 (1～65)最早的结构中，并且形成卷曲螺旋；第二个和第三个螺旋在邻近的单体上以两个相同的螺旋形成反向平行的螺旋束（Hong et al. 2008)。重要的是要指出，当尝试为结构研究生成蛋白质时，上述类型的分析很有价值，并且是选择 GAL4（1～65）用于晶体学分析的最初理由（Marmorstein et al. 1992)。

专题 10.6　KRAB 抑制结构域：个案研究

对 Krüppel 相关盒（KRAB）结构域的表征是如何分离和研究抑制结构域的极好的例子。KRAB 结构域最初被鉴定为蛋白质氨基末端的保守模体，含有这种保守模体的蛋白质带有羧基末端 *C2H2* 锌指，这种锌指结构来自通过 TFIIIA 和 Krüppel 界定的蛋白质家族［文献 Urrutia（2003）对此进行了综述］。这一模体大量存在，在大约 1/3 的 *C2H2* 基因家族均有发现，该亚家族已经被称为 KRAB-锌指蛋白（KRAB-ZFP）。保守的结构域长约 75 个氨基酸。Rauscher 及其同事证实，当将包含 4 个 KRAB-ZFP 家族成员的 KRAB 结构域的区域融合到 GAL4 DNA 结合结构域上时，它们可担当报告基因强有力的抑制因子，该报告基因在胸腺嘧啶激酶（TK）的基础启动子上游带有 5 个 GAL4 位点。TK 基础启动子带有 TATA 盒和数个反式作用因子的上游结合位点，并且因此较之含有单独 TATA 盒的启动子可产生相对较高的基础活性。

基于同源性，KRAB 结构域可分为 A 盒和 B 盒；有些家族成员有 A 盒或 B 盒，而有些家族成员两者皆有。GAL4 融合蛋白最初通过缺失分析用于研究结构域的边界（Margolin et al 1994）。在 A 盒或 B 盒中的缺失减弱了抑制作用，但缺失对 A 盒的影响更严重。的确，在这个分析中，包含 A 盒的由 45 个氨基酸组成的片段足以产生抑制作用。然后，研究人员集中对 A 盒进行了丙氨酸和基团点突变。在突变研究中，研究人员重视那些在 KRAB 结构域蛋白家族中表现出最大的保守性，并且可被鉴定为对抑制重要的特定氨基酸，对这些氨基酸残基进行突变，而不是对所有残基进行突变。这些突变体在随后的共抑制因子鉴定中是必不可少的。正如正文中所述，KRAB 结构域中的点突变对于验证与转录共抑制因子 KAP-1（Tif 1β）的相互作用是绝对必要的。通过 KRAB 结构域和 KAP1 的环 B 盒卷曲-螺旋模体，KAP-1 以低聚物形式与 KRAB 结构域相互作用。KAP-1 依次与 HP1 家族成员、组蛋白去乙酰化酶，以及一种称为 Setdb1 的新型组蛋白 3 赖氨酸 9 甲基化酶相互作用。

结构域交换的局限性

虽然结构域交换很有效，但很多研究都强调这样一个事实，即激活和 DNA 结合结构域是相当复杂的，并且依赖于细胞、启动子及位点的环境性相互作用经常调控激活结构域发挥功能的能力。而且，结构域交换有能力鉴定出蛋白质中不发挥生理学作用，但当带离环境时又能起非特异性激活结构域作用的部分。下面我们将讨论当使用结构域交换分析时必须注意的一些局限性和要考虑的事项。

首先要考虑的是一种观察结果，即激活因子/抑制因子中的某些蛋白质结构域与共激活因子/共抑制因子或其他附近的 DNA 结合蛋白相互作用，这些相互作用有助于蛋白质的 DNA 结合特异性及其在启动子特异性环境中发挥功能的能力。一个例子是淋巴增强子结合因子 1（LEF-1）蛋白。在 T 细胞受体 α（TCR-α）增强子中，LEF-1 既能弯曲 DNA，以允许其他激活因子与增强体的前后序列相互作用，同时又含有激活结构域（该激活结构域的功能尽管尚未完全了解，但依赖于 TCR-α 增强子）。因此，带有多

聚化 LEF-1 位点的简单报告基因无法发挥作用（Giese and Grosschedl 1993）。

另一种可能性是激活结构域和 DNA 结合结构域之间，以一种依赖于特定结合位点或者附近调控蛋白的方式进行联系。例如，核因子-κB（NF-κB）与数种不同的序列结合，并且似乎采取影响其活性的独特构象（Fujita et al. 1992）。的确，NF-κB 位点中的单核苷酸改变似乎影响所招募的特定共激活因子（Leung et al. 2004）。糖皮质激素受体也与应答性启动子上的数个不同位点结合，并引发不同水平的激活作用。生化实验和定点突变实验表明，DNA 结合结构域与激活结构域相互联系，以一种积极的方式控制其活性，或是将其转变为抑制因子（Lefstin et al. 1994）。这种激活结构域在结构域交换中可能不容易取得，它们的鉴定需要更先进的方法（Lefstin and Yamamoto 1998）。此外，不同的激活结构域可能依赖于环境发挥不同的功能。对糖皮质激素受体 AF1 和 AF2 激活结构域的需求依据启动子而显著不同（Rogatsky et al，2003）。

除了复杂的激活结构的活性在结构域交换中可能无法测量外，还存在另一个相反的问题，即非生理性结构域激活转录。在酵母中进行的实验中，当融合到 GAL4 DNA 结合结构域时，来自 *E. coli* 基因组的随机 DNA 序列的 5%可产生激活结构域（Ma and Ptashne 1987b）。这一观测表明在交换实验中将非生理性序列鉴定为激活结构域的可能性很大。例如，GATA-1 是一种对红细胞分化重要的转录因子，对该转录因子的结构域分析揭示，激活结构域对 GATA-1 在非红系细胞中对模式报告模板的反式激活是必要的。但是，在更多的测定 GATA-1 促进未成熟红细胞分化成终末红细胞的能力的生理分析中发现，该结构域并不是必需的（Weiss et al. 1997），相反，需要另一种细胞特异性结构域。该结果表明，最初的激活结构域虽然在交换分析中发挥功能，但是在天然环境中对激活并不是必需的。

另一个环境性问题的例子是早发性乳腺癌 1（BRCA1），BRCA1 与 DNA 修复有关。该领域的早期研究表明，BRCA1 含有一羧基端结构域，当将该结构域融合到 GAL4 DNA 结合结构域上时，可活化转录。然而，牛同源性片段的克隆显示，虽然源自奶牛的 BRCA1 羧基端结构域与人源结构域具有很好的保守性，但它含有较少的可以担当激活结构域的酸性残基（Krum et al. 2003）。牛源羧基端结构域与 GAL4 的融合会产生一种激活因子，该激活因子的激活能力远不及对应的人源结构域与 GAL4 的融合蛋白。此外，整个 BRCA1 cDNA 与 GAL4 的融合显示，BRCA1 实际上在转染分析中是一种非常差的活化因子。因此，从其天然环境中取出羧基端结构域并将其融合到 GAL4 上，揭示了一个潜在的非生理性激活结构域，该激活结构域的功能是不保守的。

这种类型的研究强调了广泛的功能分析的必要性，以及使用系统发生比较评价一个结构域是否有可能真的发挥所提议的功能。在激活结构域的情况下，这种类型的分析特别重要。抑制结构域似乎对这种人造物不太敏感，因为 *E. coli* 片段与 GAL4 的融合揭示，抑制因子的百分比（Saha et al. 1993）远比当以相似的实验鉴定激活结构时所发现的激活结构的百分比小（Ma and Ptashne 1987b）。

数量不断增加的显示位点、细胞及启动子特异性的激活/抑制结构域表明，除了在模式报告模板上进行的简单结构域交换和缺失分析外，必须采取新的方法正确地鉴定环境依赖性的结构域。这些方法包括：①在其天然应答性启动子上进行调控蛋白的初步缺

失分析；②测定缺失突变体对蛋白质整体活性的表型效应（就 GATA-1 来说是细胞分化）；③在激活因子正常活化的细胞或者密切相关的细胞类型中进行该实验。但不局限于这些方法。在模式启动子上进行结构域分析时，应确证相同的结构域在天然启动子上也是有活性的。

例如，有些研究者研究得比较深入，甚至已经在天然启动子环境中进行了结构域交换分析。在这种研究中，用 GAL4 或 LexA 位点取代启动子内的激活因子天然结合位点。将目的结构域融合到 LexA 和 GAL4 上形成嵌合性调控因子，再将这种嵌合性调控因子用于鉴定或确证环境特异性的激活结构域的功能相关性（如见 Rellahan et al. 1998）。此外，有人可能会问相反的问题，即是否可以将天然 DNA 结合结构域融合到异源激活序列上并保留功能（如见 Driever et al. 1989）。

然而，即使有了适当的启动子，某些结构域可能还需要其他细胞特异性激活因子或共激活因子的作用来发挥作用。可能仍有必要利用已经表达待研究激活因子的细胞进行共转染分析。后一方案引起一个潜在的问题，即内源激活因子可能将背景性转录增加到很高，以掩盖转染的激活因子及其突变体的效应。然而，新的方法已将转染效率提高到一个新的水平，在该水平上通常有可能将转染因子过表达到足够高，以致超过内源性水平，从而可以观察到转染因子的效应。另外，可用 RNAi 敲低内源性蛋白，并且可用一种构体替换内源性蛋白，该构体带有在密码子中的核苷酸改变，这些密码子保留了原来的氨基酸，但又不被 RNAi 打靶（Lassus et al. 2002）。此外，也可以利用密切相关但不表达该激活因子的细胞类型，或者来源于基因敲除小鼠的胚胎干细胞和胚胎成纤维细胞（MEF）来研究野生型蛋白及其突变体。这种实验会更精确地反映该蛋白质在其天然环境中的活性，同时要提醒，研究者必须切记过量表达也可能影响该分析方法的结果。例如，Smale 及其同事对 NF-κB 选择性的研究就使用了反转录病毒载体中的结构域交换，其中反转录病毒载体被插入来自 c-Rel$^{-/-}$ 小鼠的原发性单核细胞中，然后，在培养物中将原发性单核细胞分化为巨噬细胞，并监测内源性白细胞介素-12（IL-12）p40 的表达。在这些实验中，发现 c-Rel$^{-/-}$ 和 RelA（p65）的激活结构域完全可以互换。还发现 IL-12 p40的表达需要 c-Rel，因为 c-Rel 在其 DNA 结合结构域中含有 46 个氨基酸残基，这些氨基酸使 NF-κB DNA 结合结构域以高亲和性结合非一致序列 NF-κB 识别位点。也可以将 c-Rel 和 Re1A 的所有其他区域交换而不造成功能性后果（Sanjabi et al. 2000，2005）。

最后，尽管综合缺失和结构域交换方法已获得成功，但只有在极少数情况下，已经表明具有激活结构域交换的激活因子可稳定地代替天然基因产物来驱动有机体的生命周期（Driever et al. 1989，Baumann et al. 1993，1995）。因此，虽然前面所述的方法能够产生具有分析价值的结果，但重要的是不能过度解读这些实验的生理学含义。

GAL4 融合方法另外需要注意的是，它不仅适用于激活和抑制结构域的研究，同样经常用于研究其他转录辅因子。如果序列特异性 DNA 结合蛋白的作用是招募共激活因子或共抑制因子，那么很有可能在报告基因分析中，当融合到 GAL4 DNA 结合结构域上时，该共激活因子将促进转录。

细分 DNA 识别亚结构域和寡聚化亚结构域

二聚体/低聚化结构域的鉴定是蛋白质分析中的重要步骤，因为破坏二聚化界面是一种为了调控或治疗的目的使因子失活的方法。例如，Id 蛋白是缺少碱性区域的螺旋-环-螺旋（HLH）结构域（Benezra et al. 1990）。Id 及相关家族成员与碱性 HLH 蛋白质发生异源二聚化，阻断它们与 DNA 的结合，使其不能激活许多生物学过程中的转录（Norton 2000）。

有多种不同类型的二聚体模体，发现于 Jun 和 Fos 中的碱性亮氨酸拉链（bZIP）卷曲-螺旋模体（Vinson et al. 2006）和发现于 E26 中的 HLH 结构域（Massari and Murre 2000）只占其中的一小部分。的确，某些蛋白质（如 MyoD）同时包含 HLH 和亮氨酸拉链（Berkes and Tapscott 2005）。如果一个结构域属于已清楚表征的家族，如 bZIP 或 HLH，那么对它有很多了解，并且通常不需要尝试该二聚体结构域的广泛表征，除非有证据表明可能具有调控潜能的新的异二聚体对［即 Myc、Mad 及 Max 蛋白（Gallant 2006)］。的确，二聚化已知的结构和机制的使用，可以用来预测异二聚化搭档（Fassler et al. 2002；Vinson et al. 2006）。

这些模体的表征需要广泛的生化分析和分析方法，这些分析方法超出了简单的 DNA 结合分析。DNA 结合蛋白的二聚化在体外可以通过多种方法测量，包括异源二聚体分析（见第 11 章）、凝胶过滤（Li et al. 1997）、戊二醛交联（Smith et al. 1990；Wagner and Green 1993）、荧光共振能量转移（FRET）（Merickel et al. 2002）和基于凝胶的 FRET（Ramirez-Carrozzi and Kerppola 2001），以及在体内使用生物分子荧光互补（Kerppola 2006）。然而，在使用像 EMSA 一样的分析方法的 DNA 结合结构域的初步分析中，尽管二聚化亚结构域被逐渐破坏（如通过进行性缺失），通常依然能观察到特异性结合（即当比较突变型和野生型识别位点时），甚至同时观察到蛋白质对其位点的亲和力的逐渐下降，这种降低的亲和力通常是二聚体结构域已经受到损害的首要证据。

例如，即使去除二聚体亚结构域后，GAL4 和 λ 抑制因子的氨基末端片段仍然保持最小的、低亲和力的 DNA 结合能力（Pabo et al. 1979，Carey et al. 1989）。虽然在溶液中每个多肽片段都保持单体状态，截短的衍生物仍然可以通过使用残留的二聚体界面以二聚体形式结合到 DNA 上。这些通过晶体学鉴定的界面可能具有或可能不具有生理学相关性，然而它们足以驱动结合并掩盖二聚化结构域，除非仔细地测量到特定的亲和力。

如果怀疑在一个 DNA 结合蛋白的缺失分析中，已经危害到二聚体结构域，则可以进行二聚体结构域交换实验。在这种实验中，将异源二聚体结构域（即卷曲螺旋）融合到显示最小 DNA 亲和力的蛋白质上。然后，在 DNA 结合分析中比较嵌合体和非融合蛋白的亲和力。如果在 EMSA 中融合作用增强了该蛋白质的亲和力，就可以怀疑非融合蛋白质亲和力的下降是由于天然二聚体结构域已被损害或去除。为确定缺失区域的确含有二聚化结构域，则可以进行交换实验，在交换实验中将假定二聚体结构域融合到一个不同蛋白质的最小 DNA 结合结构域上。然后，分析这种互惠交换以确定其是否增强

结合。例如，Hu 和 Sauer 表明 bZIP 卷曲螺旋在体外和体内都增强 λ 抑制因子的氨基末端 DNA 结合模块的结合（Hu et al. 1990）。原理上，包含最小二聚体结构域的 GAL4（1～65）也可以与假定二聚体结构域融合。在这种情况下，亲和力的增加可能指示由假定二聚体结构域介导的增强的二聚化。

概念和策略：蛋白质-蛋白质相互作用

共激活因子和共抑制因子的分离及克隆

在讨论如何研究共激活因子和共抑制因子（辅因子）中不同的调控结构域前，我们将简要描述一下一些经典的共激活因子和共抑制因子最初是如何被鉴定并克隆的。不同的案例例证了数种通用技术，包括以抗体或相互作用蛋白探测在细菌中表达的 cDNA 文库、双杂交分析和亲和标签的使用（见专题 10. 3），以及经典的生化纯化技术。在某些情况下，共激活因子和共抑制因子产生于酵母和果蝇的遗传学研究，并且基于功能的保守概念已克隆了它们的哺乳动物对应物。

使用最久的技术之一是表达哺乳动物 cDNA 的 λ 噬菌体文库的筛选（见第 7 章）。在这些实例中，用噬菌体 λ 文库感染细菌的菌苔。将该噬菌斑转印到硝酸纤维素膜，并且使用抗目的因子的抗体或者已知与该因子相互作用的重组蛋白进行探测，结合上探针的斑点含有编码目的因子的 cDNA。例如，Eckner 及其同事通过用天然 p300 的抗体探测文库斑点，克隆了 p300 共激活因子的 cDNA，而 p300 作为一种与腺病毒 E1A 相互作用的蛋白质最初由 Harlow 发现（Harlow et al. 1986；Eckner et al. 1994）。此外，Chrivia 及其同事使用蛋白激酶 A 和 [γ-^{32}P] ATP 标记 cAMP 应答元件结合蛋白（CREB），再以 ^{32}P CREB 为探针通过探测文库分离了 CREB 的结合蛋白（CBP）（Chrivia et al. 1993）。使用类似的技术克隆了很多核受体共激活因子（综述，Glass et al. 1997）。

通过酵母双杂交分析克隆了许多蛋白质的共激活因子，其中一个典型的案例是类固醇受体共激活因子 1（SRC1）。SRC1 是在孕酮存在的条件下，通过它与孕酮受体的相互作用而克隆的（Onate et al. 1995）。双杂交分析也已用于克隆很多共抑制因子。例如，哺乳动物组蛋白脱乙酰作用复合物的 Sin3a 和 Sin3b 亚基是通过它们与 Mad 抑制结构域的相互作用而克隆的（Ayer et al. 1995）。通过双杂交还分离了 N-CoR 和 SMRT 共抑制因子，该方法鉴定了在配体不存在（而不是存在）的情况下那些结合到维甲酸和甲状腺受体上的蛋白质（综述，Shibata et al. 1997）。

生化方法已被用于鉴定众多共激活因子和共抑制因子蛋白质，之后再对这些蛋白质进行蛋白质组学分析和随后的克隆。的确，可以使用 GST 与各种配体结合结构域的融合来分离许多与共激活因子和共抑制因子相互作用的核受体（Glass et al. 1997）。类似地，通过在含有 KRAB 抑制结构域的 GST 层析柱上进行亲和层析分离了 KAP-1 共抑制因子蛋白（Friedman et al. 1996）（见专题 10. 6）。在使用 GST 维生素 D 受体（GST-VDR）从哺乳动物细胞提取物中分离调解因子复合物时，GST 亲和层析也是关键的（Rachez et al. 1998）。在这种情况下，研究人员注意到结合的蛋白质组装成一个复合

物，并使用串联质谱分析（MS/MS）和 cDNA 克隆技术来分离复合物的不同亚基；他们最初将该复合物命名为 DRIP 复合物。Roeder 在数种方法中分离了调解因子复合物。在其中一种方法中，他使用了 FLAG 标记的甲状腺受体，从细胞提取物中分离了甲状腺激素相关蛋白（TRAP）复合物（Fondell et al. 1996）。在其他方法中，还可使用生化分析来鉴定体外刺激激活因子依赖性转录的成分（Ito and Roeder 2001）。Tjian 实验室也使用功能分析方法来分离调解因子复合物，并且之后使用 MS 和 cDNA 克隆技术来分离单个亚基（Naar et al. 1999）。

总之，大量基于亲和力或活性的实验技术已用来分离共激活因子和共抑制因子。在这些技术只能鉴定单个亚基的情况下，接着使用其他方法来鉴定其余的亚基。例如，可用另外的双杂交分析或在哺乳动物细胞中表位标记亚基的表达（见专题 10.3）来鉴定共激活因子和共抑制因子的相互作用亚基。必须强调的是，这些辅因子的克隆还需要一些测试以验证结合是特异的。例如，辅因子与以上鉴定的激活结构域或抑制结构域结合，但不与发生突变的结构域结合。另外，已表明共激活因子和共抑制因子以依赖于激活/抑制信号（即 CREB 磷酸化对于 CBP 或核受体的配体依赖性对于 SRC1）的方式与激活因子和抑制因子结合。在下面的几节中，我们将更详细地讨论如何进一步表征这些相互作用。

研究激活结构域与共激活因子相互作用的方法

确定激活结构域相互作用的搭档对于理解该结构域如何调控转录是共同而重要的目的。与 Pol II 或哺乳动物 TFIID 中亚基之间的强力相互作用不同，激活因子与靶位点间的相互作用相对微弱。因此，确定它们的相互作用是一项极富挑战性的工作。已有多种研究方法可供研究者使用，其中某些方法在前文中已进行了讨论，但在这里还将进一步详细地说明。一个激活区域的结构-功能分析的目的是，使用突变体作为工具鉴定或确证激活因子的蛋白质靶点。有数种用于鉴定和表征在大分子复合物环境外的单个蛋白质-蛋白质相互作用的生化和遗传学方法，这些方法包括特异性配基亲和层析柱的使用、免疫共沉淀、化学交联及酵母双杂交分析。另外的一些分析方法，如荧光各向异性和表面等离子体共振（SPR），非常专业，并且可用于细化该分析。

有两种哲学上的方法：①模型引导法，在这种方法中激活结构域被选择是由于它与可能靶点的相互作用；② 随机法，在这种方法中使用遗传学或生化技术从潜在靶点的集合中鉴定假定靶位。模型引导法假设一个步骤是重要的，并受激活因子的影响，如激活因子通过影响 TFIID 的结合来影响转录，因为 TFIID 的结合是复合物组装的最早的步骤。然后，通过以完整的激活/抑制结构域进行亲和层析测试这一观点，并使用激活/抑制结构域的突变体证实这一观点。尽管其相对简单，但模型引导法或许成为一种危险的起点，因为研究者可能猜测一种蛋白质-蛋白质的相互作用很重要。这样的研究可产生一种模型，该模型很难经受得住随后对其可靠性的检验。随机法更谨慎，往往判断力更强。在这个方法中，一旦确定了某个结构域的重要性，则使用双杂交分析、遗传检验，以及粗提取物或大量纯化因子的亲和层析等做广泛尝试，以鉴定其靶点。我们讨论简单亲和树脂的使用和这些方法的缺陷，因为大多数的注意事项对更复杂的方法也是适

用的。重要的是要强调方法的复杂性不一定是其有效性的度量。表面等离子体共振和亲和层析一样，易于提供异常的结果，虽然其结果更加量化。

一个需要考虑的关键问题是激活因子和“靶”蛋白（即共激活因子或通用机器）的相互作用可能发生在只有当激活因子、通用因子和共激活因子组装成转录复合物时才显示出来的表面上。在该复合物环境中，这些相互作用可能是高度协同地接触网络的一部分。然而，在单独的（即孤立的）情况下，这些相互作用可能非常弱。如果该相互作用确实很弱，就使得这种相互作用难以测量并且难以证实其生理相关性。由于这种原因，最后必须进行功能性生化研究，从而将在简单的双组分结合分析中观察到的相互作用和对体内外转录及转录复合物组装的影响相互联系起来。我们将在下面和第 12 章中详尽说明这一点。

通过亲和层析研究激活/抑制结构域与其靶标之间的相互作用

原理

历史上，亲和层析已经被证实是用于鉴定相互作用的蛋白质搭档的最为强大的生化技术之一（图 10.5）。该技术已被广泛用于研究激活/抑制结构域与其靶点间蛋白质-蛋白质的相互作用。鉴定转录因子搭档的经典方法之一是由 Greenblatt 及其同事发展起来的，该研究小组将 RNA PolⅡ共价连接到树脂上，并且使用这种产生的亲和基质从粗提物中分离 RAP30 和 RAP74。值得注意的是，RAP30 和 RAP74 被证明是 TFⅡF 的小亚基和大亚基（Sopta et al. 1985）。这种方法非常灵活，并且在该领域的早期，Green 及其同事研究 VP16 激活结构域与通用转录机器相互作用时就采用了这种方法（Lin and Green 1991；Lin et al. 1991）。综上所述，亲和层析已被证明是一种用于鉴定辅因子的非常有用的工具。

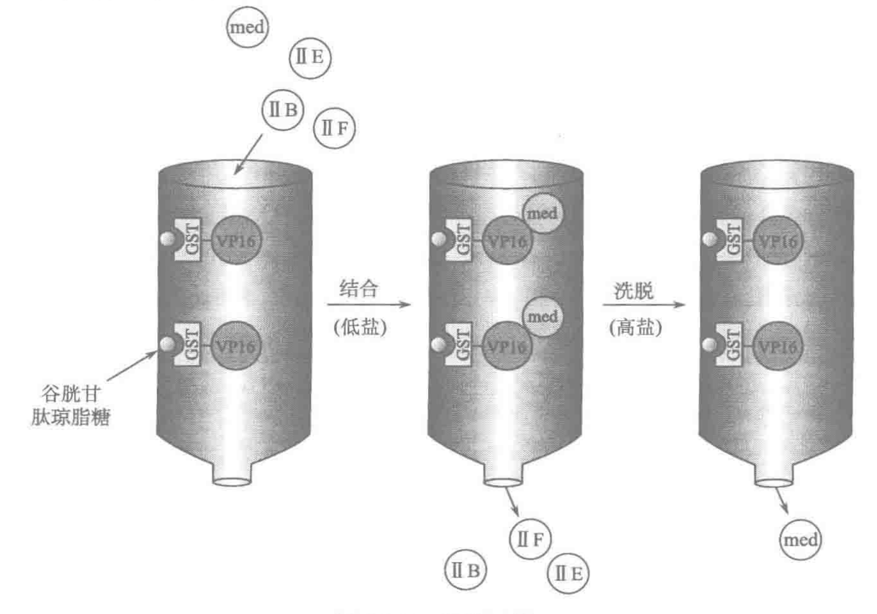


图 10.5 亲和层析。

亲和层析有多种变体，在基因调控领域中使用最为普遍的两种方法包括：使用通过一种蛋白质与基质的共价连接产生的树脂，以及使用 GST 融合蛋白结合的谷胱甘肽 Sepharose/琼脂糖。后者由于其相对简单而受到越来越多的欢迎。与商业来源的树脂共价交联，是一种相当简单的产生亲和层析柱的技术，这种技术的例子包括 CNBr 活化的 Sepharose（GE Healthcare 公司，cat. #17-0820-01）或者 Affi-Gel 10 和 15 树脂（Bio-Rad）。Affi-Gel 中含有 *N*-羟基琥珀酰亚胺活化的基质，这种基质可与出现在赖氨酸中或蛋白质的氨基末端的伯胺交联。缓冲液中必须不含伯胺以防止交联效率的降低（例如，HEPES 是一种在交联应用中常用的缓冲液，而 Tris、羟胺，以及乙醇胺不合适，因为它们会和化学交联基团发生反应)。通常，要交联的纯化蛋白质以每毫升加入的树脂中 1～20 mg 蛋白质的密度与激活的树脂孵育。混合物轻摇 1 h 至过夜使其发生交联，然后，未结合的蛋白质用缓冲液洗涤去除。之后，通过加入乙醇胺中和亲和性基质，并且用结合缓冲液预平衡，为亲和性结合分析做好准备。加入保护剂（如 EDTA 或叠氮化物）可长期冷冻或冷藏保存。某些层析基质不能承受反复冻融，反复冻融会使微珠崩解，因此，这类基质应避免冷冻。

重要的是要改变基质上配体的密度并检查其对靶蛋白结合的效果。已经表明，配体密度对结合的蛋白质的量和种类都有显著影响（见 Greenblatt and Ingles 1996）。在某些情况下，较低密度的基质不能与靶蛋白结合，而较高密度的基质可能开始表现出显著的非特异性的相互作用。因为已经清楚地证实了许多原核及真核蛋白质能够非特异性地与琼脂糖及纤维素基质相结合，所以在这些研究中也有必要制备对照层析柱，对照层析柱中含有单独的树脂或与突变型蛋白质相连接的树脂，这些对照亦应在开始分析前准备好。

与其将蛋白质与树脂共价结合，不如将蛋白质与纯化标签相融合（见第 11 章），这些纯化标签如 GST、麦芽糖结合蛋白（MBP）或 His6（见专题 10.2）。生成的融合蛋白再分别连接到相应的亲和性基质上，如谷胱甘肽-琼脂糖、淀粉/Sepharose 或 Ni-NTA/Sepharose。其中，GST 融合方法在基因调控结构域可能应用得最广泛。

一旦获得合适的基质，就可以将其与假定的靶蛋白或粗蛋白混合液（即核提取物）在不同的结合条件下混合。基质经缓冲液洗涤以除去未结合的蛋白质，然后，用高浓度盐溶液、配体或特制的缓冲液洗脱结合的蛋白质。之后，可将结合和未结合的馏分通过 SDS-聚丙烯酰胺凝胶电泳（SDS - PAGE)、免疫印迹及 MS 进行检验。另外，可在试管中分批进行结合反应，在孵育和（或）洗涤后，使用低速离心以沉淀微珠。

亲和方法的警告

虽然亲和方法简单而诱人，但它也困难重重。在亲和分析中，对于特异性的正确对照是比较野生型和突变型激活因子/抑制因子对于一个假定靶蛋白或蛋白质混合物的结合。然而，对于真核生物激活结构域，我们并不总是完全了解激活结构域突变体的性质，同时，在没有结构信息的情况下，我们也不确信突变是否仅特异性地影响对蛋白质-蛋白质相互作用重要的氨基酸，或者突变是否对结构具有整体影响。尽管如此，该方法已被证实在多种情况下非常有效。两个杰出的例子是两个不同的实验室各自从粗提物中分离了调解因子（mediator）复合物，这些研究代表了上一节中所描述的随机法的

例子。在一个案例中，使用 GST-VDR 分离最初被命名为 DRIP 的蛋白质复合物，该复合物在维生素 D 存在的条件下结合到层析柱上，反之，则不结合（Rachez et al. 1998）。如上所述，该复合物被证实是调节因子。在另一个案例中，腺病毒 E1A 锌指激活结构域 CR3 与调节因子结合，然而一个不能支持激活的突变体却不结合（Boyer et al. 1999）。因此，在前一个研究中，使用配体依赖性验证相互作用的相关性，而在后一个研究中，使用的是激活结构域的突变体。

过度解释亲和分析会产生危害，例子之一是从 20 世纪 90 年代早期主要集中于 TBP 与激活因子相互作用的大量研究，这些研究是上述模型引导法的例子。由于 TBP 是最早被克隆的通用转录因子的一个亚基，并且可广泛获得，因此很多激活因子因与 TBP 相互作用而被研究，却没有考虑转录复合物中几十个其他的多肽，这些多肽可能代表更好的候选者。野生型激活因子被证实与 TBP 相互作用，然而缺失体或点突变体被证实在亲和结合分析中不与 TBP 相互作用。在许多出版物中，单凭这些数据作为该相互作用有效性的唯一证据，但是，随后的实验却表明某些激活结构域与 TBP 的底面上的碱性氨基酸残基相互作用，这种相互作用应该是抑制，而不是促进 TBP 与 DNA 的结合。然而，在许多情况下，由于未检测到激活因子对 TBP 结合到 TATA 盒上的影响，所以从未确定该相互作用的生化后果。在其他情况下，后续的研究清楚地显示，激活结构域“黏附”许多不同的转录组分，使得最初的观察更不肯定。在大多数研究中，这些相互作用仍然没有被整理出来，因为这些研究从来没有超越该分析的起始阶段。除了忽视了相互作用对转录复合物组装的生化后果外，几乎没有什么研究曾阐明在 TBP 中是否存在互惠突变体，这些突变体能废除转录分析中与该激活因子的相互作用及对该激活因子的应答（如见 Tanseyand Herr 1995）。这样的分析会发现靶蛋白上重要的表面，并且有助于确定这些相互作用从生理角度上看是否有意义。

总之，应始终使用互惠突变方法和生物化学方法支持最初的亲和力观测。很少设法解决这样的基本问题，但考虑到围绕哺乳动物转录中蛋白质-蛋白质相互作用的不确定性，这些基本问题对于确证一个相互作用的相关性很关键。

改变特异性遗传系统

在第 8 章中为 DNA 结合蛋白质所讨论的改变特异性方法是一个杰出的案例，在该案例中用巧妙的实验设计分析结构域的生物学功能。该方法使研究人员能够剖析在必然表达内源性该蛋白质的细胞内蛋白质-蛋白质相互作用的相关性。如果不改变特异性，内源性蛋白质就会干扰该分析。改变特异性的基本前提与遗传学中等位基因特异性抑制的概念相似。在等位基因特异性抑制中，一个组分中的突变被另一个组分中的特异性补偿突变所抑制。原则上，补偿突变对最初的突变具有特异性，并且发挥作用以修复两个相互作用并互补的氨基酸之间或一个氨基酸与一个碱基对之间的直接接触。实际上，很少能观察到完美的等位基因特异性抑制蛋白（suppressor），但是由于某些抑制蛋白仍具有一定的等位基因特异性，并能改变相互作用的特异性，因此等位基因特异性抑制蛋白的概念仍是有效的。等位基因特异性抑制蛋白方法已用于构建具有新的序列特异性的 DNA 结合蛋白，并用于了解蛋白质-蛋白质相互作用。迄今为止，这种方法在真核转录

系统中已经应用于数种 DNA 结合蛋白，包括 TBP 和 TFIIB 的少数通用转录机器组分，以及数个共激活因子和共抑制因子（讨论如下）。

TBP

在 TBP 存在的情况下，DNA 结合表面上某些氨基酸的变化会改变 TBP（TBP_{AS}）的特异性，使它以高亲和力识别 TGTAAA（Strubin and Struhl 1992)。这种相互作用对于野生型 TBP 来说是被禁止的，原因是外环小沟鸟嘌呤上的亚氨基与 TBP 的 DNA 结合表面之间的空间冲突。突变的氨基酸明显形成一个与亚氨基相适应的口袋。尽管这种突变体也结合野生型 TATAAA 盒，但它结合 TGTAAA 盒的能力意味着可以创建在体内只对 TBP_{AS}应答的人工启动子。在这样一种系统中，可将在 TGTAAA 上游带有激活因子结合位点的报告模板与 TBP_{AS}及待测激活因子共转染。因为只有 TBP_{AS}与 TGTAAA 结合，才能研究激活因子与 TBP_{AS}之间的相互作用（图 10.6A)。虽然该系统最初发展于酵母中，但在哺乳动物细胞中也能很好地发挥作用。例如，该系统已用于分析在体内与通用转录机器相互作用的 TBP 的表面（Bryant et al. 1996)，下面我们将会讨论这一话题。

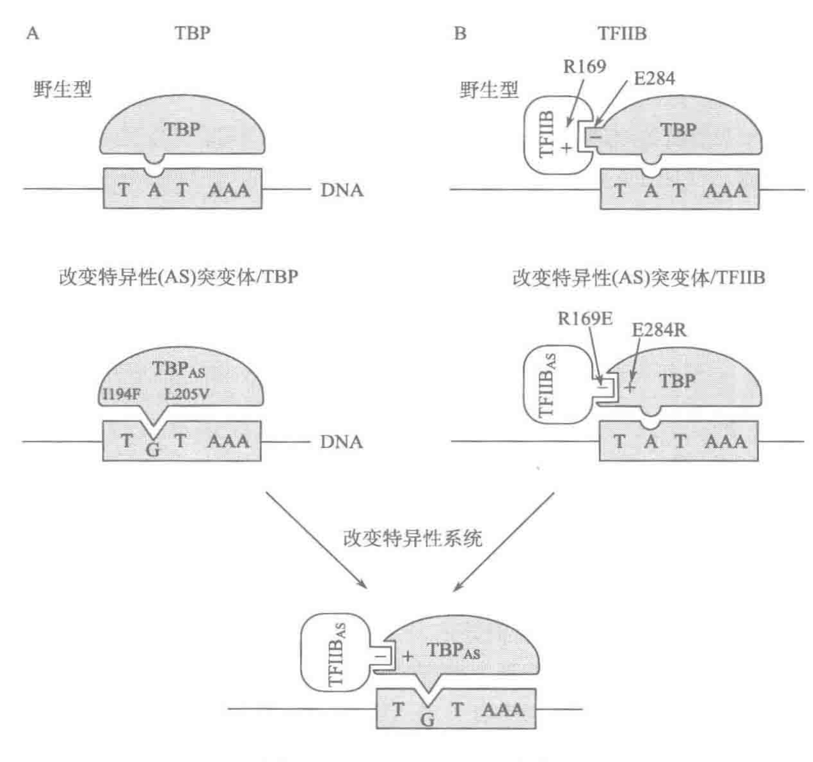


图 10.6　改变特异性突变体。

TFIIB

目前已经构建了改变特异性的 TFIIB，这种 TFIIB 可以与 TBP_{AS}相结合来研究激活因子与 TFIIB 的相互作用（图 10.6B）（Tansey and Herr 1997)。在 TATA/TBP/

TFIIB复合物的晶体结构中，TBP 中第 284 位上的 Glu 与 TFIIB 中第 169 位的 Arg 相互作用。TBP 中第 284 位的 Glu 突变为 Arg 则废除这种相互作用。但是，TFIIB 中 Arg 突变为 Glu 的抑制性突变又恢复了该相互作用。该系统可用于研究转录激活过程中不同激活因子如何与 TFIIB 相互作用。

FOG-1

改变特异性方法也适用于共激活因子和共抑制因子。的确，酵母双杂交筛选和体内哺乳动物细胞分化分析的联合使用被证实是一个强有力的组合，可用于破解被称为“Friend of GATA-1”，或 FOG-1 的共激活因子蛋白参与红细胞分化的生理作用（Cantor and Orkin 2005）。如上所述，GATA-1 是一种控制红细胞和巨核细胞成熟的激活因子。GATA-1 是经酵母双杂交筛选分离的含有两个锌指的转录因子家族的一部分，其羧基末端的锌指与 DNA 结合，而氨基末端的锌指与共激活因子 FOG-1 相互作用。FOG-1 也是一种锌指蛋白，它通过自己的 6 号指与 GATA-1 相互作用。GATA-1 和 FOG-1 在多种细胞类型中共表达，但是很难确定它们在红细胞分化中是具有协同的作用还是独立的作用，因为敲除的表型差异很小。

为了阐明 GATA-1 和 FOG-1 在体内的作用，Orkin 及其同事发展了改变特异性系统（Crispino et al. 1999）。研究人员使用酵母双杂交方法来分离 GATA-1 氨基末端锌指的突变体，这些突变体废除了 GATA-1 与 FOG-1 的 5 号和 6 号指的相互作用。研究人员避免了会影响锌指折叠的半胱氨酸突变，通过 EMSA 确证突变的 GATA-1 可结合到 DNA 上，但在缺少 GATA-1 的 G1E 前体细胞中不能诱导红细胞分化。在 GATA-1 中获得了突变体后，他们突变 FOG-1 以获得能够在酵母双杂交分析中与 GATA-1 突变体结合的补偿性突变体。结果分离到一个在 6 号指中发生突变的 FOG-1，该突变体适度结合 GATA-1 突变体。通过用免疫共沉淀法从转染细胞提取物中验证了该相互作用后，研究人员创建了带有 FOG-1 和 GATA-1 突变体的稳定的 G1E 细胞系。将 GATA-1 突变体融合到雌激素受体（ER）配体结合结构域上。使用 ER 配体结合结构域（LBD）融合的原因是在配体缺失的条件下，LBD 抑制了许多被连接蛋白的功能。而添加配体可逆转这种抑制作用，并允许融合蛋白发挥功能。一旦加入雌激素来激活突变体 GATA-1，在突变体 FOG-1 存在的情况下，显著比例的细胞发生了分化，而当加入野生型 FOG-1 时，只有少部分细胞发生了分化。该数据说明激活因子 GATA-1 与共激活因子 FOG-1 之间的特异的相互作用对细胞分化确实至关重要。在文献 Crispino 和 Orkin（2002）中可找到该方法的全面描述和注意事项。

改变特异性的方法需要注意的是，维持结构而废除相互作用的点突变通常不会施加绝对的影响。相反，它们反映出两个相互作用蛋白质的亲和力的下降。因此，搭档蛋白质的过表达可以降低部分甚至是全部的由突变造成的影响。正如在 FOG-1 的情况下，野生型蛋白依然与突变体 GATA-1 相互作用，虽然这种相互作用较弱，但它在过表达时对红细胞系的分化产生了一些营救作用。然而，由于仔细并定量地进行了这些实验，差异是明显的。如果研究人员在该生物分析中没有观察到差别，他们可能会断定 FOG-1 与 GATA-1 的相互作用并不是绝对必要的，或者他们可能会尝试降低活化因子和共激活因子的表达水平以接近它们在细胞中的自然生理浓度。

Mad

改变特异性的方法也已用于研究共抑制因子（Cowley et al. 2004）。Mad 蛋白是一个 DNA 结合碱性螺旋-环-螺旋-亮氨酸拉链（bHLHZ）蛋白质家族；该家族蛋白质与 Max 异二聚化，并通过与含有哺乳动物 Sin3a 亚基（mSin3a）的共抑制因子相互作用而抑制转录。Mad 蛋白都含有 32 个氨基酸的氨基末端结构域，该结构域与 mSin3a 和 mSin3b 相互作用，因而称为 Sin3 相互作用结构域（SID）。SID 与成对的 mSin3a 和 ni-Sin3b 的双亲性螺旋 2 结构域（PAH2）相互作用。PAH2 和 SID 之间的复合物的结构已经确定（Sheppard et al. 2001）。Eisenman 及其同事以此结构为指导，通过产生各种补偿性疏水氨基酸的变化，能改变 PAH2 和 SID 的疏水性界面。其中，补偿性疏水氨基酸的变化废除了突变体与野生型结构域的表面互补性，但允许两组突变体在各种结合分析中相互之间能发生相互作用（Cowley et al. 2004）。重要的是，当建成完整的 mSin3A 时，通过突变体 SID，PAH2 突变体在体内恢复了抑制，尽管抑制效果不如野生型蛋白那么显著。功能部分恢复背后的原因尚不清楚。研究人员推测，mSin3A 是很多多亚基抑制复合物的部件，也许因为与内源性蛋白的竞争作用，转染的 mSin3A 突变体没有完全组装成抑制复合物。的确，在用任何多亚基复合物尝试这种类型的分析时，这个问题就变得非常相关，因为简单的转染分析可能未允许突变体完全装配成复合物。此外，当构建稳定的细胞系时，这些突变体可能产生有害影响。

显然，生物环境和所使用的分析方法的类型将对实验的可行性具有显著的影响，并且在实施一个长期的项目之前应加以考虑。总之，为了研究转录机器中蛋白质-蛋白质间的相互作用，有很多产生改变特异性突变体的方法。可用基于结构的方法做特定的预测，或者可用双杂交分析为补偿性改变提供一个基本的选择系统。在开始分析前，必须考虑用于分析改变特异性系统功能的分析方法的生物学特性和可行性。

通用转录机器的结构-功能分析

哺乳动物通用转录机器的突变分析一般在生化系统中进行，这些生化系统以体外转录和对其他通用因子的亲和力为读出，如上所述，其原因是通用转录机器在所有细胞中都具有功能，内源性通用因子很有可能会干扰任何在体内进行突变分析的尝试。缺失分析最初用于鉴定不同通用因子的功能所必需的结构域。虽然对于某些蛋白质（如 TFIIB）来说，这种方法被证明易于驾驭（Ha et al. 1993），但对于其他蛋白质来说，其结果难以解释，并且会将注意力集中在点突变体的使用上。

然而，当蛋白质结构已知，并且可用突变发生回答与机制相关的非常具体的问题时，定点突变的使用才是最令人兴奋和最能说明问题的。迄今为止，研究得最好的例子是 TBP 的突变研究，设计这些突变研究的目的是用于了解 TBP 的哪些表面参与同其他通用因子的相互作用（图 10.7）（Bryant et al. 1996，Tang et al. 1996），该方法将可能被推广用于激活因子与共激活因子相互作用的研究。

以 TBP 晶体结构为指导（综述见 Burley and Roeder 1996），Ebright 及其同事将丙氨酸扫描与 EMSA 分析相结合来鉴定在体外无法与通用转录因子相互作用的人 TBP 突变体（Tang et al. 1996）。总体策略是仅对 DNA-蛋白质复合物中能进入到溶剂中的氨

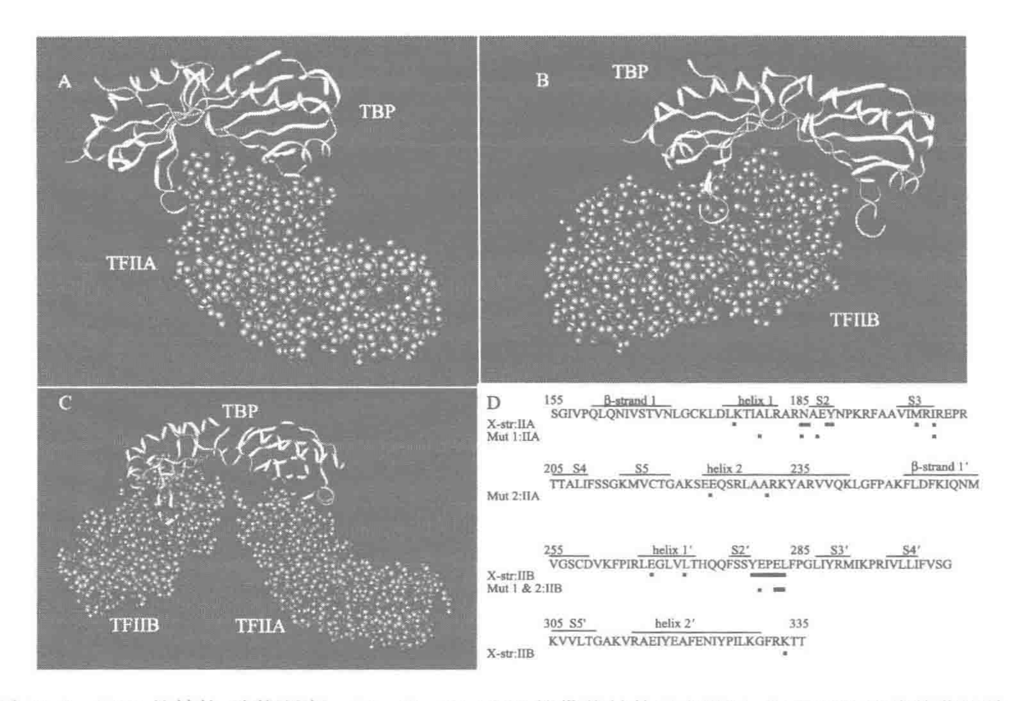


图 10.7　TBP 的结构-功能研究。(A、B、C) TBP 的带状结构及 TFIIA 和 TFIIB 的范德华渲染。为清楚起见，图上省略了 DNA。在晶体结构中显示与 TFIIA（A、C）及 TFIIB（B、C）相互作用的 TBP 残基用黑色球棍表示。箭头指出了在正文中已描述的突变研究中影响 TFIIA 和 TFIIB 结合的 TBP 残基。TBP 的序列（D）采用 Nikolov 等（1996）的编号方案。该结构的综述见 Nikolov 和 Burley（1997）。序列上方以直线形式标出了 TBP 晶体结构中螺旋 1 和 2、螺旋 1′ 和 2′，以及折叠片 1～5 和折叠片 1′～5′ 的位置。以粗体标示的残基在晶体中与 DNA 接触。X-strIIA 是指 TBP 中带有粗体下划线的残基，这些残基与 TBP、TFIIA 及 DNA 的酵母共复合物中的 TFIIA 相互作用（Tan et al. 1996）。这些接触被叠加到人 TBP 结构中的同源位置上。X-strTFIIB 是指在 TBP、TFIIB 及 TATA 的晶体结构中与 TFIIB 相互作用的 TBP 残基（Tan et al. 1996）。突变 1 说明了通过基因置换涉及 TFIIA 和 TFIIB 结合的残基（Bryant et al. 1996）。突变 2 说明了通过丙氨酸置换涉及 TFIIA 和 TFIIB 结合的残基。通过突变分析鉴定到了依靠晶体学分析鉴定的 TFIIA 和 TFIIB 界面中的残基，其中突变 2 例外，突变 2 涉及螺旋 2 中的残基，螺旋 2 据称与 TFIIA 相互作用，但晶体结构中未显示这一点。有两点可能解释这一冲突，即人和酵母 TBP 之间的差异，以及对于螺旋 2 相互作用重要的残基从酵母结构（不是所有的 TFIIA 都被结晶）中丢失的可能性［晶体结构图例由 Michael Haykinso（UCLA）利用分子图像结构建模计算机程序洞察者 II（Molecular Graphics structure modeling computer program InsightII）制作］。

基酸残基使用丙氨酸扫描。DNA-结合残基及参与结构包装的残基不发生突变，突变体最终包括 TBP 上的 81 个非脯氨酸表面残基。研究者筛选了那些导致通用因子对 TBP-TATA 复合物亲和力下降 5 倍以上的表型。他们计算出这相当于 1kcal 的能量，这一能量被宣称对应一个普通的侧链相互作用。该研究的目的是鉴定对 TBP 与 TATA 盒的结合没有影响，但却破坏 EMSA 实验中其他通用因子结合的突变体。

然而不同突变体对 TFIIA、TFIIB、TFIIF 及 Pol II 的结合有明显的影响，而它们

对 TFIIE 和 TFIIH 结合的影响微乎其微。通过将突变的位置叠加到 TBP 表面上，这项研究可以为这些因子如何组装到 TBP 以形成基础转录复合物建立模型。虽然这项研究是在 TBP-TFIIA 和 TBP-TFIIB 及 DNA 的三元复合物晶体结构确定之前进行的，但是由突变发生推断出的 TBP-TFIIB 界面与通过晶体学方法鉴定的 TBP 上的 TFIIB 结合位点非常吻合（综述见 Burley and Roeder 1996）。相比之下，从研究所推断的 TBP 上的 TFIIA 结合位点和随后的 TATA/TBP/TFIIA 三元复合物的晶体结构不相符合（Tan et al . 1996）。

另一个杰出的实例中，Berk 及其同事进行了人源 TBP 综合性结构-功能研究（Bryant et al. 1996）；在该研究中，利用基团置换分析了 TBP 的 89 个表面可接触的残基。研究中，不仅将突变发生与 TFIIA 和 TFIIB 的体外结合研究相结合，还将突变发生与细胞培养物的转染分析相结合，在转染分析中使用 TBP_{AS}应答性启动子测量被激活的转录。结果发现，在先前已经鉴定的 TBP 上的 TFIIB 界面中的基团置换能降低基础的和激活的转录水平，该结果与共晶体结构及 Ebright 等进行的独立研究相一致（Tang et al. 1996）。然而，Berk 的研究也鉴定了参与 TFIIA 结合的氨基酸残基，这与已发表的晶体结构非常吻合。在 TBP 的表面上鉴定到了另外两个的突变体簇，这些突变体簇代表了一些潜在结合位点，它们可能是对激活转录重要的蛋白质的结合位点，这些蛋白质可能是 TBP 相关因子（TAF）或其他未被鉴定的辅因子。尚不清楚为什么一项研究正确地鉴定了 TFIIA 结合位点（Bryant et al. 1996），而另一项研究则不能（Tang et al. 1996）。两个重要的差别包括：①一项研究使用基团置换突变体，这种突变体能活跃地破坏界面，而另一项研究使用丙氨酸扫描，这种技术被认为只是导致接触的丧失；② 一项研究将体外分析和功能性体内转录相结合，这可能有助于解决结合实验中的含糊之处。

总之，有多种多样的突变方法和功能分析方法。所有定点突变研究的确认并不仅仅依赖于功能的丧失，还依赖于证明突变体影响非常明确的生化过程，但同时需使蛋白质的总体结构保持完整。

共激活因子的分析

如上所述，共激活因子和共抑制因子的分析可能是成问题的，因为很多这类因子是多亚基复合物。然而，基因表达的许多多样性是由作为激活因子和抑制因子与共激活因子和共抑制因子独特相互作用的结果而产生的。因此，考察这些蛋白质是如何被分析的一些案例，有助于在开始新的分析时为使用的技术和考虑的策略提供指导。我们将讨论两个例子，其一是 p300/CBP，该蛋白质主要以单一亚基存在，而另一个是调解因子，是多亚基复合物。

p300/CBP

p300 最初是由 Harlow 鉴定的，作为数种 E1A 互作蛋白中的一种，这些 E1A 互作蛋白是通过免疫共沉淀法从腺病毒感染的细胞提取液中分离的（Harlow et al. 1986）。该蛋白质与 CBP 密切相关；两者都是约 2400 个残基的蛋白质，并具有 61%的一致性。如早先所述，编码 p300 的基因由 Eckner 在 1994 年从 293 细胞的 cDNA λ 噬菌体表达

文库中克隆（Eckner et al. 1994），并显示含有溴域，溴域的功能当时是未知的，但是它的存在却是数个已知共激活因子的一个特征（Mujtaba et al. 2007）。Eckner 以及其他人提出 p300 是一个转录共激活因子，它介导激活因子对基因转录的影响。CBP 由 Chrivia 克隆，他使用^{32}P 标记的 CREB 探测了 λ 表达文库（Chrivia et al. 1993）。小鼠中 p300 和 CBP 的基因敲除是胚胎致死性的。各种研究已经证实，p300 和 CBP 共享某些重叠的功能，但就绝大部分而言，它们似乎在发育与分化中发挥不同的作用（见 Kasper et al. 2006，以及其中的参考文献）。1996 年，Nakatani 及其同事证明 p300 和 CBP 含有固有的 HAT 活性（Ogryzko et al. 1996）。除了其 HAT 活性外，p300 还有具有将自身和像 p53 一样的其他转录因子乙酰化的能力（Vo and Goodman 2001）。文献中经常提到的两个共激活因子就是指 p300/CBP 或 CBP/p300，但它们是不同的蛋白质。

界定 p300/CBP 功能的最强大的工具之一是缺失分析。使用限制酶和其他方法删除该基因的区域，并且之后对蛋白质进行体外结合研究和细胞转染分析，引人注目的是，发现 p300/CBP 惊人地经得起缺失。同样地，缺失方法被用于定位很多相互作用搭档的结合位点。已知或被强烈怀疑相互作用的蛋白质的当前的数目已超过 300 个（http://www.stjude.org/brindle）。p300/CBP 通过至少 5 种不同的表面与激活因子相互作用、与其他共激活因子［如 SRC-1 和 p300/CBP 相关因子（PCAF）］相互作用，以及与通用转录机器的很多部件相互作用（Vo and Goodman 2001）。

这些蛋白质的各种不同的研究已经使用了不同的技术来研究它们与不同的激活因子的相互作用。我们不能广泛地涵盖全部这些技术，某些技术已在其他章节中涉及，然而，我们将描述如何在体外证实某些特定的相互作用，并在细胞中进行验证。用于分析 p300 和 CBP 相互作用的经典范例在 20 世纪 90 年代早期产生于 Montminy 和 Goodman 实验室［在文献 Vo and Goodman（2001）中做了综述］。这些研究人员当时正在研究 CREB 激活因子刺激对环腺苷酸（cAMP）应答的转录的机制。已知 CREB 通过蛋白激酶 A（PKA）对 cAMP 信号作出应答，激活像生长抑素一样的基因（Montminy et al. 1996）；PKA 磷酸化位于一个 CREB 激活结构域内的 Ser133。CREB 包含一个氨基末端富含 Gln 残基的结构域及带有 Ser133 的激酶可诱导结构域（KID），这两种结构域都是使用上面几节中所讨论的 GAL4 融合方法鉴定的。如前所述，CBP 特异性地与在 Ser133 上被磷酸化的 CREB 结合。

可用多种技术定位 CBP 与 CREB 的 KID 相互作用的表面（即 KIX 结构域）。这些技术包括分析 KID 与不同 p300/CBP 缺失/点突变体间蛋白质-蛋白质相互作用的体外方法［即 GAL4 融合、远 Western 印迹（far western blot）、免疫共沉淀、亲和层析及 EMSA 分析］。此外，可用多种体内技术验证该相互作用，包括激活因子及 p300/CBP 突变体与应答性启动子的共转染，以及显性失活方法。事实上，在体内研究的早期，使用 CBP 中和抗体的显微注射证实了 CREB 介导的转录激活对 CBP 的需要（Montminy et al. 1996），而 KIX 结构域对转录的显性失活效应可用于支持该结构域在细胞中与 CREB 相互作用的想法（Parker et al. 1996）。

在此处及在其他激活因子的相关研究中，重要的一点是将结合研究与功能分析联系起来，并用 CBP 和激活结构域的突变体验证结果。在 KIX-KID 案例中，研究者幸运地

发现两者间的相互作用依赖于Ser133磷酸化，因为Ser133突变是一个抑制CREB-KID或GAL4-KID融合蛋白激活潜力的简单方法。此外，因为KIX结构域小，从CBP删除它一般不影响该蛋白质的其他活性。例如，在许多情况下，当融合到GAL4上时，CBP的缺失突变体仍然可以刺激转录。的确，KIX结构域在进化中似乎具有保守性，并且在其他蛋白质（如Med15调解因子亚基）中亦有发现（Novatchkova and Eisenhaber，2004）。目前，已经知道了数个具有激活结构域的KIX结构域的结构（见Yang et al. 2006，及其中的参考文献）。

最近，用于分析p300/CBP的技术已经包括基因敲除动物，以及来自这些小鼠的胚胎干细胞（ES）和胚胎成纤维细胞（MEF）的使用。p300/CBP在小鼠中的敲除是胚胎致死性的，但已构建了条件性敲除，以研究p300/CBP在某些组织中的功能（Kasper et al. 2006），并且已建立了ES和MEF细胞系（见Roth et al. 2003）。同样，也使用了基因敲入方法。虽然p300和CBP突变体在HAT结构域中的敲入是胚胎致死性的，但是子宫中胚胎的详细表型和ES细胞的特性是完全不同的（如Shikama et al. 2003）。因此，除了一系列用于分析p300/CBP在细胞培养物中过表达的分析方法外，从基因敲除和敲入研究中已经获取很多信息。在不能获得p300敲除（即肿瘤细胞中p300的分析）的研究中，RNAi技术已成功用于敲低p300（Iyer et al. 2007）。

调解因子

激活因子和抑制因子与多亚基的共激活因子和共抑制因子复合物的相互作用的研究非常复杂。一个亚基的不同结构域或区域的缺失可能不仅影响与抑制因子或激活因子的相互作用，也影响各亚基间的相互作用或者复合物的催化活性。此外，这些体内相互作用的研究受到这样一种事实的妨碍，即用于研究许多激活因子和抑制因子及与p300/CBP类似的单亚基共激活因子的瞬时转染分析，在用于分析带有多亚基的复合物时，更有挑战性。瞬时转染分析可能不会给亚基留下足够的时间，使其组装成适当的复合物，并且如果突变的共激活因子扮演显性失活的角色，稳定细胞系的建立也可能会有问题。某些共激活因子可以通过转染基因敲除的ES或MEF细胞系，或者采用RNAi干扰实验进行研究，但这些方法都有一定的难度。可按很多研究p300的方法体外研究共激活因子（Kraus and Kadonaga 1998；Kraus et al. 1999；Manning et al. 2001），但同样也具有诸多的局限性。

由于这样的原因，该领域的许多研究者已选择在一种体外结合分析简单表明激活因子或抑制因子与共激活因子或共抑制因子的相互作用，并且之后使用RNAi技术敲低该亚基，并废除该相互作用对内源性基因或者对报告基因的影响。此外，可用ChIP实验证明相互作用的因子共定居到内源基因上。尽管这种类型的分析是有意义的，但它未阐明特定相互作用的生理学相关性，因为RNAi和敲除细胞受到间接影响且ChIP是一个关联性技术。因此，改变特异性突变体的使用可能是必要的，以便为该相互作用具有相关性提供最强有力的支持。即使在使用改变特异性方法的情况下，也必须准备好一种强大的分析方法以评价该相互作用的生物学特性。

为了讨论特定的DNA结合蛋白如何通过蛋白质-蛋白质的相互作用招募多亚基共激活因子并与它们相互作用，我们将涉及甲状腺受体（TR）与包含在MED1

(TRAP220) 内的 LXXLL 模体的相互作用，MED1 是核心调解因子复合物中最大的亚基（见第 1 章和第 12 章）。TR 是核受体超家族的成员，如上所述，它以配体依赖性的方式激活转录。这个家族的大部分成员含有位于配体结合结构域内的转录激活结构域，该转录激活结构域被称为 AF2。已知 AF2 通过被称为 LXXLL 模体或核受体盒的模体，结合到如 SRC1 的共激活因子或其他因子上。例如，许多核受体与 p300 的 LXXLL 模体相互作用。注意可以通过蛋白质序列的分析来定位 Med1 内的核受体盒；蛋白质序列的分析可使用用于鉴定蛋白质中功能性位点的真核线性模体（Eukaryotic Linear Motif, ELM）资源（http://elm.eu.org/）。通过在 HeLa 细胞中表达 FLAG 标记的 TR，Roeder 及其同事们能够分离出一种蛋白质复合物，该复合物在配体三碘甲状腺乙酸或 T3 的存在下与 TR 结合。他们将该复合物称为 TRAP，表示 TR 相关蛋白（Fondell et al. 1996）。尽管现在已知调解因子是由 30 多个亚基组成的，但最初的纯化仅揭示出 10 个亚基（见第 1 章和第 12 章）。

为了鉴定哪个亚基将 TR 特异地拴到了调解因子上，Roeder 及其同事进行了 far Western 分析。在这个分析中，将调解因子在 SDS 凝胶中分级分离并印迹到膜上。之后，在 T3 存在和缺失的情况下，以通过体外翻译产生的 ^{35}S 标记 TR 探测该印迹膜（见专题 10.2）。研究发现，在 T3 存在的情况下，TR 特异性地与印迹上的一个 220 kDa 的蛋白质结合。研究人员已经克隆了调解因子 10 个亚基的不同 cDNA，并且能够在免疫共沉淀和 GST 亲和分析中使用体外翻译的蛋白质表明在 T3 存在时 Med1 亚基（最初称为 TRAP220）特异性地结合到 TR 上。的确，通过构建 GST 与其他许多核受体的融合蛋白，发现 Med1 结合到许多核受体家族成员上，从而表明通过一个共同的机制，一个家族的相关蛋白质可将调解因子招募到一个基因上。有趣的是，Med1 包含两个 LXXLL 模体，如上所述，已知 LXXLL 模体与核受体配体结合结构域相互作用（Yuan et al. 1998）。因此，这一观察结果直接表明 TR 与 Med1 相互作用的潜在表面。为了阐明这一表面的相关性，研究人员在该模体中生成了缺失和点突变。

在该分析的这个阶段，技术变得相当复杂。研究人员首先将 Med1 的野生型和突变型 LXXLL 的 cDNA 转染进入细胞。然而，他们仅仅观察到野生型 Med1（并非突变型）对 TR 应答性的报告基因产生微弱的刺激作用（Yuan et al. 1998）。有人提出该刺激作用微弱是因为细胞系中存在内源性调解因子，并且在标准转染条件和时线下转染的 Med1 并没有完全组装成有转录活性的调节因子复合物。为了单独证明 LXXLL 与 TR 的相互作用，研究者使用了显性失活的方法。当转染仅含有 LXXLL 模体的 Med1 部分时，该蛋白质片段强烈地抑制 TR 介导的转录。

当获得了基因敲除小鼠时，进一步验证该相互作用成为可能（Ito et al. 2000）。基因敲除的小鼠在早期妊娠阶段发生死亡，但基因敲除的 MEF 细胞可用于研究 TR 功能的基本方面。在共转染分析中，TR 显示出对 TR 报告基因具有微弱但可检测到的激活作用，这种微弱的激活作用可通过共转染 Medl 表达载体加以营救（Ito and Roeder 2001）。如上所述的显性失活方法，其结果为 TR 的确通过 Med1 发挥功能提供了额外的支持。基因敲除 MEF 细胞的主要优点是它们允许对功能重要的特定 Med1 结构域进行分析。首要地，从基因敲除的 MEF 中分离出相对完整的缺少 MED1 的调解因子，从

而表明 MED1 对该复合物的稳定性不是至关重要的。的确，除 TR 和核受体超家族成员外，缺少 Med1 的调解因子在体外和体内能很好地对多种激活因子作出应答。然后，才有可能使用 MEF 分析 Med1 缺失和点突变，结果表明，带有两个 LXXLL 模体的氨基末端 670 个氨基酸的部分组装成调解因子复合物，该复合物在体内外对 TR 和 T3 作出应答并发挥功能。在 MEF 细胞中，突变任意一个 LXXLL 模体仅废除与 TR 的相互作用，必须同时突变两者才能废除 TR 应答性报告基因的转录激活作用。的确，后来的研究表明，不同的 LXXLL 模体特异地与不同的核受体相互作用（Ren et al. 2000）。这项研究阐明了数个重要的特征。第一，当 Med1 亚基被删除时，共激活因子本身仍具有功能并对其他激活因子作出响应。第二，MEF 细胞可用于分析 Med1 表面上与 TR 相互作用的特定突变。第三，单个 LXXLL 模体中的突变废除了体外结合，但要废除体内 TR 的应答性则要求两个 LXXLL 模体的突变。第四，这些突变体 Med1 亚基组装成了有功能的调解因子复合物。

MEF 细胞的使用从生化和生物的角度来看对验证生理相关性至关重要。但 MEF 不可能对所有目前的共激活因子都有效，并且研究者可能希望在不同的细胞中研究一种相互作用，以了解这种相互作用在如癌症中的作用。在这种情况下，可用 RNAi 从细胞中去除一个亚基。事实上，RNAi 已成功地用于从乳腺癌细胞中缺失 Med1 亚基，以证明 Med1 亚基对 ER 介导的转录的必要性（Zhang et al. 2005）。在 RNAi 处理的细胞系中难以进行结构-功能分析，因为它会阻碍敲低蛋白突变体的加入。在原理上和实践中，如果 cDNA 含有改变了 RNAi 结合位点但仍保持密码子的点突变［即将甘氨酸密码子 CCG 改变为 CCC（Lassus et al. 2002）］，那么就可以将突变体亚基引入 RNAi 处理细胞。但是，这种方法是棘手的，并且需要通过密码子中的改变，敲低内源性的 RNA、不影响转染的 RNA，并且不影响表达的蛋白质。所有生物表现出一定的密码子偏爱，并且替代密码子的使用有可能在某些情况下造成不利的后果。

技术

方案 10.1　PCR 介导的定点突变

与传统的定点突变不同，这种方法使用完整质粒扩增来产生点突变体，仅需要单一的 PCR 步骤（图 10.8）（Weiner et al. 1994）。该方法可将小的突变引入启动子位点并且更适合于将单突变或双突变引入蛋白质。Stratagene 公司出售一种与下述方案相似的试剂盒（QuikChange 点突变试剂盒，200518 和 200519），该方法简单易行，而且任何使用标准蛋白质表达载体的实验室以商业获得的试剂都能很容易使用该技术。

合成一对通常长 30～35 个核苷酸的编码目的突变体的引物，突变体两侧各有 15 bp 的同源编码序列。设计的错配不应多于 3～4 个核苷酸（2 个核苷酸的错配最容易掺入；包含多于 2 个核苷酸错配的引物效率较低），并且应位于引物的中间部位。注意，在突变发生期间产生或去除限制性位点可简化产生克隆的分析。而且，每条引物的解链温度（T_m）必须等于或高于 78℃，以便该方法更好地进行，这通常意味着引物必须含

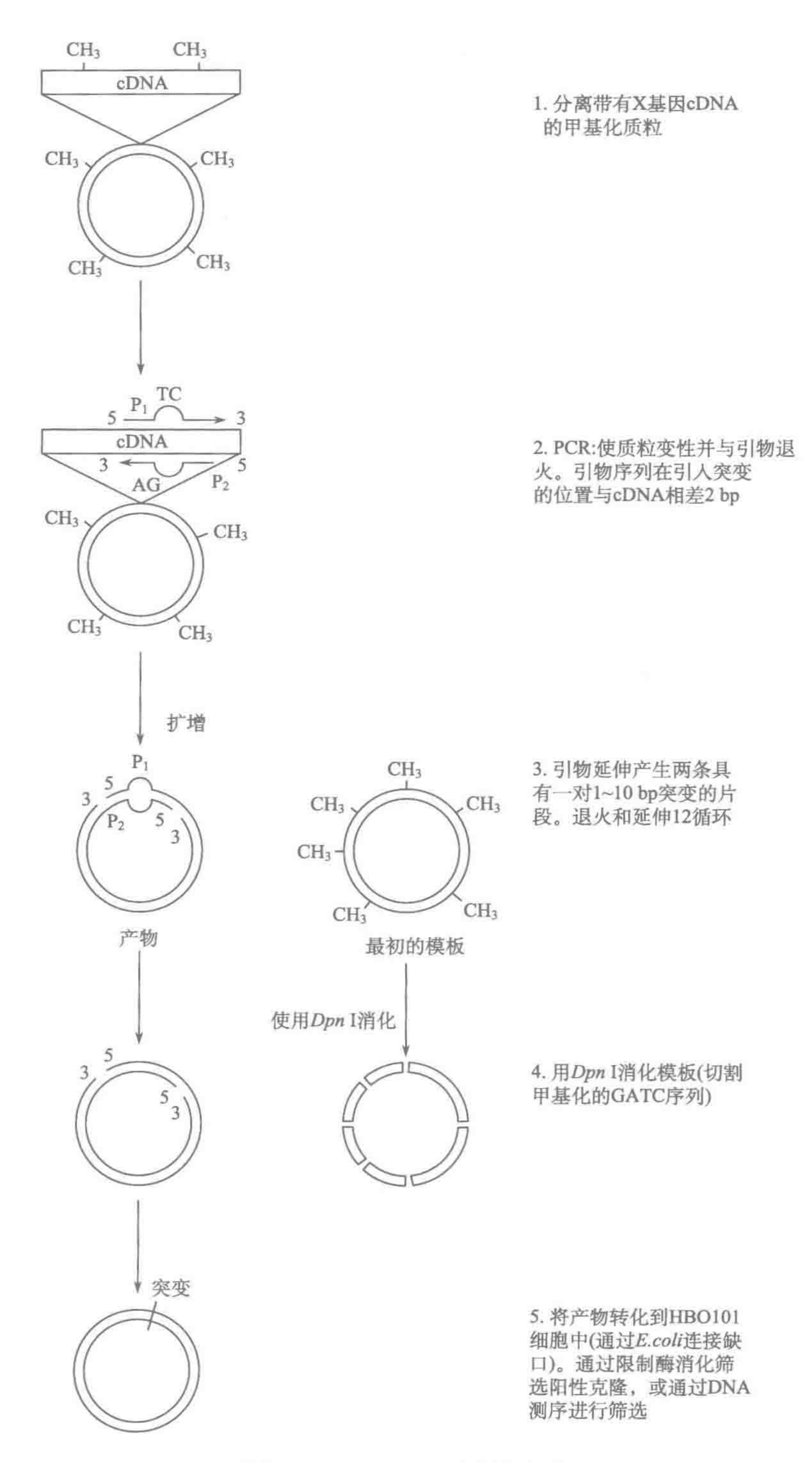


图 10.8　PCR 介导的定点突变。

有至少50%的G-C。有些商业获得的程序可确定为某一区域设计的引物的最佳 T_m 值；或者，使用下述公式计算给定引物的 T_m 值：

$$T_m = 81.5 + 0.41(\%\ \text{G-C}) - 675/N - \%\ 错配$$

其中，N 为以碱基对为单位的引物长度。

同时，在DNA腺嘌呤甲基转移酶阳性（dam^+）*E. coli* 菌株（如HB101）中生长带有目的区域的质粒（来自大多数常用菌株中的DNA都被甲基化；如果对用于扩增预成质粒的菌株没有把握，应在进行突变前通过 *Dpn*I 检查其切割）。使用CsCl密度梯度法或商业获得的试剂盒分离质粒DNA。将引物加入质粒中并且在热循环仪中用Pfu DNA聚合酶扩增。使用Pfu聚合酶代替 *Taq* 酶是因为其校对功能和随机突变的可能性减少。Pfu聚合酶绕质粒一周，产生一系列互补的全长单链线性DNA分子，这些DNA分子起始于引物并且终止于引物5′端前一个核苷酸。两个互补的线性分子杂交以形成含有突变位点的完整质粒环。杂交体在每条引物5′端之前也含有两个切刻（nick）。30 bp左右的互补性引物足以使质粒环化并转化 *E. coli*。重组杂交体未被甲基化，然而来自 *E. coli* dam^+ 菌株的亲本DNA在某些限制性位点上被高度甲基化。对于12个循环来说，将产生最高产量为原始数量12倍的质粒（注意：扩增不是指数式的）。因此，产物在凝胶上的分析和检测可能并不会显示出明显的扩增。

然后，用 *Dpn*I 切割反应产物；*Dpn*I 仅在5′-G^{m6}ATC-3′序列处切割dam-甲基化的DNA，这可以有效地从混合物中去除甲基化的亲本DNA。用未被切割的子代突变体（sibling mutant）分子转化 *E. coli* 并且筛选。实验中加上不含引物的模拟反应非常关键，以确保 *Dpn*I 彻底去除亲本分子。还有必要确保在延伸阶段引物的5′端和质粒稳定地杂交，否则Pfu聚合酶会置换引物，从而产生能转化 *E. coli* 的亲本分子。通常，5′端离突变15 bp对引物是足够的，但是引入额外的核苷酸会增加突变体生产的效率。通过由突变发生引入或去除限制性位点，可轻松地筛选从小量制备分子中获得的产物，以鉴定突变体。当不能引入合适的位点时，必须进行测序分析来鉴定重组分子。一般情况下至少有20%的产物发生了突变，并且通常有90%～100%的产物发生了突变。

时线与安排

突变发生很简单，并且不需要任何特殊的操作。一般来说，获得引物后，可在一天之内完成扩增、亲本分子的消化，以及随后的转化。第二天进行培养克隆并制备重组质粒。通过突变添加或去除限制性位点可以促进筛选，否则就必须进行测序。正如所有的PCR突变方法，进行二重反应获得可用的备份克隆是明智的，以防PCR过程向质粒引入其他的突变。由于具有校对功能的热稳定性DNA聚合酶（如Pfu）的使用，重复反应变得越来越少见。

第1天：突变发生

第2天：质粒制备和限制性分析

第3天：序列分析

提纲

步骤1：使用PCR产生突变体质粒DNA（1 h，30 min）。

步骤 2：使用 *Dpn*I 切割亲本 DNA（1 h，30 min）。

步骤 3：转化合适的 *E. coli* 宿主菌（1～2 h 及培养过夜）。

步骤 4：质粒制备（小量制备）和限制性分析（1 天及过夜）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

琼脂糖凝胶（1%）

dNTP 混合物（25 mmol/L）

*Dpn*I

E. coli，dam$^+$ 菌株（如 HB101）

如果用于筛选阳性克隆的限制酶对甲基化不敏感，则将质粒转化入 dam$^+$ 的 *E. coli* 菌株。

溴化乙锭（EB）〈!〉

HB101 感受态细胞

矿物油

Pfu DNA 聚合酶（2.5U/μl，Stratagene 600135）

10×Pfu 反应缓冲液

质粒（100 ng/μl）

通过 CsCl 密度梯度离心法或商业获得的层析柱（如 QIAGEN，27104）分离自 dam$^+$ *E. coli* 菌株（如 HB101）。

引物（100 pmol/μl）

基于计算机分析和促进筛选的限制性位点的可得性，订购或合成合适的引物。使用前需通过快速蛋白质液相色谱或凝胶电泳进行纯化。

设备

Eppendorf 管（硅化的，0.6 ml；或同等物）

凝胶电泳仪

加热器，预设为 65℃

质粒小量制备试剂盒（QIAGEN）

热循环仪

紫外灯〈!〉

水浴锅，预设为 37℃

程序

步骤 1：使用 PCR 产生突变体质粒 DNA

做二重反应，并且做一个缺少引物的模拟反应。这将对后来的步骤中确定 *Dpn*I 对

亲本 DNA 的切割效率很有必要。

1. 冰上操作，在 0.6 ml 的 Eppendorf 管或专门的 PCR 管中准备如下反应。

10×Pfu 反应缓冲液	5.00μl
引物 1（100 pmol/μl）	1.25μl
引物 2（100 pmol/μl）	1.25μl
dH_2O	39.50μl
含有 cDNA 的质粒（100 ng/μl）	1.00μl
pfu DNA 聚合酶（2.5U/μl）	1.00μl
dNTP 混合液（25 mmol/L）	1.00μl
总体积	50.00μl

注意当使用高度感受态的 *E. coli* 细胞时，质粒 DNA 和引物的量可降低 10 倍。

2. 反应混合液用矿物油覆盖（除非使用具有隔热室的 PCR 仪）。
3. 在热循环仪上进行线性扩增：

	变性	退火	延伸
预变性	96℃，30 s		
12 循环	96℃，30 s	55℃，1 min	68℃，6～12 min（每 kb 质粒至少 1 min）

重要的是延伸温度要适当地低于引物的 T_m 值。

步骤 2：以 *DpnI* 切割 *PCR* 产物

1. 如果必要，可将反应混合液转移至新管中。加入 1～5U 的 *Dpn*I。
2. 37℃水浴 1～2 h。
3. 消化后，将反应管在 65℃加热 10 min。

在转化 *E. coli* 之前不需要进行酚抽提和沉淀。

步骤 3：转化合适的 *E. coli* 宿主

1. 将 45 μl 限制性内切核酸酶切割的混合液加入到 100 μl HB101 感受态细胞中。

由于突变发生的性质，在某些情况下转化效率可能非常低。小心不要将上层的矿物油带进转化体系中，因为这会抑制转化效率。

2. 包括三个转化对照实验：

- 用于检测细胞感受态的超螺旋质粒。
- 用于检测 *Dpn*I 消化效率的模拟反应（如缺乏引物或 Pfu 聚合酶）。
- 用于确保细胞独自不能在所用选择性抗生素中生长的无质粒对照。

步骤 4：产物分析

1. 分析信噪比。

带有以模拟反应产物转化的细胞的平板上大量克隆的出现，表明 *Dpn*I 切割无效。

在这种情况下，任何下一步的操作将是不明智的。此外，以突变发生混合液转化的平板上，克隆数增加 2～3 倍，则可保证继续筛选。某些试剂盒以含有 *lacZ* 的突变质粒作为对照，在这种情况下，要包括一个可修复 *lacZ* 功能的引物，并且反应效率可通过 lac-*E. coli* 菌落的蓝白斑筛选进行分析。

2. 进行质粒的少量制备（QIAGEN）。

3. 用合适的限制酶消化 DNA 以筛选由于突变作用引起的限制性位点的丧失或获得。

4. 在含有溴化乙锭的 1%琼脂糖凝胶上分析限制性产物，用紫外灯使产物可见。

附加说明

筛选重组体

很明显，不可能总是存在与要引入的突变相一致的限制性位点，特别是在进行综合性丙氨酸扫描分析时（详见本章）。我们发现在短时间内，并且通过遵照引物 T_m 规则，产生于特定突变的克隆几乎全部呈该突变阳性。这允许将该技术大规模用于产生定点突变体，但突变体最后必须通过 DNA 测序加以确证。

对 Dam 甲基化敏感的限制性位点

有些酶不能切割生长在 dam^+ 菌株中的质粒上的甲基化 DNA 位点。例如，如果 DNA 被 dam 甲基化酶甲基化在腺嘌呤 N^6 位置，*Bcl*I 限制酶就不能识别 *Bcl*I 位点。许多常见的限制酶（全部名单见供应商的目录）都不能识别甲基化的腺嘌呤。将突变的质粒转化进 dam 缺陷型的 *E. coli* 菌株中，允许通过 dam 甲基化敏感的酶对质粒 DNA 进行随后的筛选。然而，关键是要在 dam^+ 菌株中生长欲突变的质粒 DNA。

多重突变

在许多现行的商业化试剂盒中，可加入多重引物，以便在同一基因的不同位置上同步引入突变。在这种情况下，扩增反应会产生数个延伸产物，这些延伸产物都与同一个模板杂交但被切刻分开。试剂盒同时含有热稳定 DNA 连接酶，该酶可以将切刻连接在一起，产生一个完整的、可转化到 *E. coli* 中的 DNA 分子。

疑难解答

仅获得亲本质粒

当从突变转化和模拟转化中比较性地筛选菌落时，观察到了良好的信噪比，但筛选时，重新获得亲本质粒。

可能原因：这明显是由引物 T_m 低引起的。可能 T_m 太低时，Pfu 绕质粒一周后会置换掉引物。

解决办法：重新合成具有正确 T_m 的新引物。

无转化子

可能原因：细胞的转化效率低。

解决办法：使用亲本质粒检查细胞的转化效率。

可能原因：如果转化效率没问题，这种结果可能是由亲本 DNA 的扩增效率低造成的，其中有数个可能的原因。其一是扩增循环数太少，其二是大质粒的扩增效率低，其三因为突变的大小超过 2～3 个核苷酸，该程序的效率剧烈下降，正如使用许多突变方法的情况。

解决办法：增加循环和重复次数，并且增加延伸时间。如果一切都失败了，则通过 PCR 连接引入突变并亚克隆到 cDNA 的适当区域中。

参考文献

Amann E.，Brosius J.，and Ptashne M. 1983. Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in *Escherichia coli*. *Gene* 25：167-178.

Ayer D. E.，Lawrence Q. A.，and Eisenman R. N. 1995. Mad-Max transcriptional repression is mediated by ternary complex formation with mammalian homologs of yeast repressor Sin3. *Cell* 80：767-776.

Baumann R.，Warren G.，and Askovic S. 1995. Restoration of the Epstein-Barr virus ZEBRA protein' s capacity to disrupt latency by the addition of heterologous activation regions. *Virology* 211：64-72.

Baumann R.，Grogan E.，Ptashne M.，and Miller G. 1993. Changing Epstein-Barr viral ZEBRA protein into a more powerful activator enhances its capacity to disrupt latency. *Proc. Natl. Acad. Sci.* 90：4436-4440.

Benezra R.，Davis R. L.，Lockshon D.，Turner D. L.，and Weintraub H. 1990. The protein Id：A negative regulator of helix-loop-helix DNA binding proteins. *Cell* 61：49-59.

Berkes C. A. and Tapscott S. J. 2005. MyoD and the transcriptional control of myogenesis. *Semin. Cell Dev. Biol.* 16：585-595.

Boyer T. G.，Martin M. E.，Lees E.，Ricciardi R. E，and Berk A. J. 1999. Mammalian Srb/Mediator complex is targeted by adenovirus E1A protein. *Nature* 399：276-279.

Brent R. and Ptashne M. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* 43：729-736.

Bryant G. O.，Martel L. S.，Burley S. K.，and Berk A. J. 1996. Radical mutations reveal TATA-box binding protein surfaces required for activated transcription in vivo. *Genes Dev.* 10：249 1-2504.

Burley S. K. and Roeder R. D. 1996. Biochemistry and structural biology of transcription factor IID (TF II D). *Annu. Rev. Biochem.* 65：769-799.

Cantor A. B. and Orkin S. H. 2005. Coregulation of GATA factors by the Friend of GATA (FOG) family of multitype zinc finger proteins. *Sernin. Cell Dev. Biol.* 16：117-128.

Carey M.，Leatherwood J.，and Ptashne M. 1990. A potent GAL4 derivative activates transcription at a distance in vitro. *Science* 247：710-712.

Carey M.，Kakidani H.，Leatherwood J.，Mostashari F.，and Ptashne M. 1989. An amino-terminal fragment of GAL4 binds DNA as a dimer. *J. Mol. Biol.* 209：423-432.

Chrivia J. C.，Kwok R. P.，Lamb N.，Hagiwara M.，Montminy M. R.，and Goodman R. H. 1993. Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* 365：855-859.

Cowley S. M.，Kang R. S.，Frangioni J. V.，Yada J. J.，DeGrand A. M.，Radhakrishnan I.，and Eisenman R. N. 2004. Functional analysis of the Mad1-mSin3A repressor-corepressor interaction reveals determinants of specificity，affinity，and transcriptional response. *Mol. Cell. Biol.* 24：2698-2709.

Cress W. D. and Triezenberg S. J. 1991. Critical structural elements of the VP16 transcriptional activation domain. *Science* 251: 87-90.

Crispino J. D. and Orkin S. H. 2002. The use of altered specificity mutants to probe specific protein-protein interactions involved in the activation of GATA-1 target genes. *Methods* 26: 84-92.

Crispino J. D., Lodish M. B., MacKay J. P., and Orkin S. H. 1999. Use of altered specificity mutants to probe a specific protein-protein interaction in differentiation: The GATA-1: FOG complex. *Mol. Cell* 3: 219-228.

Cunningham B. C. and Wells J. A. 1989. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* 244: 1081-1085.

Davis R. L., Cheng P. F., Lassar A. B., and Weintraub H. 1990. The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* 60: 733-746.

de Boer H. A., Comstock L. J., and Vasser M. 1983. The tac promoter: A functional hybrid derived from the trp and lac promoters. *Proc. Natl. Acad. Sci.* 80: 21-25.

de Boer E., Rodriguez P., Bonte E., Krijgsveld J., Katsantoni E., Heck A., Grosveld F., and Strouboulis J. 2003. Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proc. Natl. Acad. Sci.* 100: 7480-7485.

Driever W., Ma J., Nusslein-Volhard C., and Ptashne M. 1989. Rescue of bicoid mutant Drosophila embryos by bicoid fusion proteins containing heterologous activating sequences. *Nature* 342: 149-154.

Eckner R., Ewen M. E., Newsome D., Gerdes M., DeCaprio J. A., Lawrence J. B., and Livingston D. M. 1994. Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes Dev.* 8: 869-884.

Emami K. H. and Carey M. 1992. A synergistic increase in potency of a multimerized VP16 transcriptional activation domain. *EMBO J.* 11: 5005-5012.

Fassler J., Landsman D., Acharya A., Moll J. R., Bonovich M., and Vinson C. 2002. B-ZIP proteins encoded by the *Drosophila* genome: Evaluation of potential dimerization partners. *Genome Res.* 12: 1190-1200.

Flint J. and Shenk T. 1997. Viral transactivating proteins. *Annu. Rev. Genet.* 31: 177-212.

Fondell J. D., Ge H., and Roeder R. G. 1996. Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proc. Natl. Acad. Sci.* 93: 8329-8333.

Friedman J. R., Fredericks W. J., Jensen D. E., Speicher D. W., Huang X. P., Neilson E. G., and Rauscher F. J., 3rd. 1996. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev.* 10: 2067-2078.

Fujita T., Nolan G. P., Ghosh S., and Baltimore D. 1992. Independent modes of transcriptional activation by the p50 and p65 subunits of NF-κB. *Genes Dev.* 6: 775-787.

Gallant P. 2006. Myc/Max/Mad in invertebrates: The evolution of the Max network. *Curr. Top. Microbiol. Immunol.* 302: 235-253.

Giese K. and Grosschedl R. 1993. LEF-1 contains an activation domain that stimulates transcription only in a specific context of factor-binding sites. *EMBO J.* 12: 4667-4676.

Gill G. and Ptashne M. 1987. Mutants of GAL4 protein altered in an activation function. *Cell* 51: 121-126.

Giniger E. and Ptashne M. 1987. Transcription in yeast activated by a putative amphipathic α helix linked to a DNA binding unit. *Nature* 330: 670-672.

Giniger E., Varnum S. M., and Ptashne M. 1985. Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell* 40: 767-774.

Glass C. K., Rose D. W., and Rosenfeld M. G. 1997. Nuclear receptor coactivators. *Curr. Opin. Cell Biol.* 9: 222-232.

Gnatt A. L., Cramer P., Fu J., Bushnell D. A., and Kornberg R. D. 2001. Structural basis of transcription: An RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292: 1876-1882.

Goodrich J. A. , Hoey T. , Thut C. J. , Admon A. , and Tjian R. 1993. *Drosophila* TAF II 40 interacts with both a VP 16 activation domain and the basal transcription factor TF II B. *Cell* 75: 519-530.

Greenblatt J. and Ingles C. J. 1996. Interaction between acidic transcriptional activation domains of herpes simplex virus activator protein VP16 and transcriptional initiation factor II D. *Methods Enzymol*. 274: 120-133.

Greil F. , Moorman C. , and van Steensel B. 2006. DamID: Mapping of in vivo protein-genome interactions using tethered DNA adenine methyltransferase. *Methods Enzymol*. 410: 342-359.

Gstaiger M. , Georgiev O. , van Leeuwen H. , van der Vliet P. , and Schaffner W. 1996. The B cell coactivator Bob1 shows DNA sequence-dependent complex formation with Oct-1/Oct-2 factors, leading to differential promoter activation. *EMBO J*. 15: 2781-2790.

Ha I. , Roberts S. , Maldonado E. , Sun X. , Kim L. U. , Green M. , and Reinberg D. 1993. Multiple functional domains of human transcription factor II B: Distinct interactions with two general transcription factors and RNA polymerase II. *Genes Dev*. 7: 1021-1032.

Hagmann M. , Georgiev O. , and Schaffner W. 1997. The VP16 paradox: Herpes simplex virus VP16 contains a long-range activation domain but within the natural multiprotein complex activates only from promoter-proximal positions. *J. Virol*. 71: 5952-5962.

Harlow E. , Whyte P. , Franza B. R. , Jr. , and Schley C. 1986. Association of adenovirus early-region 1A proteins with cellular polypeptides. *Mol. Cell. Biol*. 6: 1579-1589.

Herr W. 1998. The herpes simplex virus VP16-induced complex: Mechanisms of combinatorial transcriptional regulation. *Cold Spring Harbor Sytnp. Quant. Biol*. 63: 599-607.

Hong M. , Fitzgerald M. X. , Harper S. , Luo C. , Speicher D. W. , and Marmorstein R. 2008. Structural basis for dimerization in DNA recognition by Gal4. *Structure* 16: 1019-1026.

Hope I. A. and Struhl K. 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* 46: 885-894.

Hu J. C. , O' Shea E. K. , Kim P. S. , and Sauer R. T. 1990. Sequence requirements for coiled-coils: Analysis with λ repressor-GCN4 leucine zipper fusions. *Science* 250: 1400-1403.

Ito M. and Roeder R. G. 2001. The TRAP/SMCC/Mediator complex and thyroid hormone receptor function. *Trends Endocrinol. Metab*. 12: 127-134.

Ito M. , Yuan C. X. , Okano H. J. , Darnell R. B. , and Roeder R. G. 2000. Involvement of the TRAP22O component of the TRAP/SMCC coactivator complex in embryonic development and thyroid hormone action. *Mol. Cell* 5: 683-693.

Iyer N. G. , Xian J. , Chin S. F. , Bannister A. J. , Daigo Y. , Aparicio S. , Kouzarides T. , and Caldas C. 2007. p300 is required for orderly Gl/S transition in human cancer cells. *Oncogene* 26: 21-29.

Johnston M. 1987. A model fungal gene regulatory mechanism: The GAL genes of *Saccharomyces cerevisiae*. *Microbiol. Rev*. 51: 458-476.

Kasper L. H. , Fukuyama T. , Biesen M. A. , Boussouar F. , Tong C. , de Pauw A. , Murray P. J. , van Deursen J. M. , and Brindle P. K. 2006. Conditional knockout mice reveal distinct functions for the global transcriptional coactivators CBP and p300 in T-cell development. *Mol. Cell. Biol*. 26: 789-809.

Keegan L. , Gill G. , and Ptashne M. 1986. Separation of DNA binding from the transcription- activating function of a eukaryotic regulatory protein. *Science* 231: 699-704.

Kerppola T. K. 2006. Visualization of molecular interactions by fluorescence complementation. *Nat. Rev. Mol. Cell Biol*. 7: 449-456.

Kraus W. L. and Kadonaga J. T. 1998. p300 and estrogen receptor cooperatively activate transcription via differential enhancement of initiation and reinitiation. *Genes Dev*. 12: 331-342.

Kraus W. L. , Manning E. T. , and Kadonaga J. T. 1999. Biochemical analysis of distinct activation functions in p300 that enhance transcription initiation with chromatin templates. *Mol. Cell. Biol*. 19: 8123-8135.

Kristie T. M. and Sharp P. A. 1990. Interactions of the Oct-1 POU sub-domains with specific DNA sequences and with the HSV α- transactivator protein. *Genes Dev.* 4: 2383-2396.

Krum S. A., Womack J. E., and Lane T. F. 2003. Bovine BRCAI shows classic responses to genotoxic Stress but low in vitro transcriptional activation activity. *Oncogene* 22: 6032-6044.

Lai H. C., Qian B., and Tu C. P. 1989. Characterization of a variant rat glutathione 5-transferase by eDNA expression in *Escherichia coli*. *Arch. Biochem. Biophys.* 273: 423-432.

Lassus P., Rodriguez J., and Lazebnik Y. 2002. Confirming specificity of RNAi in mammalian cells. *Sci. STKE* 2002: PL13.

Lefstin J. A. and Yamamoto K. R. 1998. Allosteric effects of DNA on transcriptional regulators. *Nature* 392: 885-888.

Lefstin J. A., Thomas J. R., and Yamamoto K. R. 1994. Influence of a steroid receptor DNA-binding domain on transcriptional regulatory functions. *Genes Dev.* 8: 2842-2856.

Leung T. H., Hoffmann A., and Baltimore D. 2004. One nucleotide in a κB site can determine cofactor specificity for NF-κB dimers. *Cell* 118: 453-464.

Li X., Lopez-Guisa J. M., Ninan N., Weiner E. J., Rauscher F. J., 3rd, and Marmorstein R. 1997. Overexpression, purification, characterization, and crystallization of the BTB/POZ domain from the PLZF oncoprotein. *J. Biol. Chem.* 272: 27324-27329.

Lin Y. S. and Green M. R. 1991. Mechanism of action of an acidic transcriptional activator in vitro. *Cell* 64: 971-981.

Lin Y. S., Ha I., Maldonado E., Reinberg D., and Green M. R. 1991. Binding of general transcription factor TF II B to an acidic activating region. *Nature* 353: 569-571.

Luo Y. and Roeder R. G. 1995. Cloning, functional characterization, and mechanism of action of the B-cell-specific transcriptional coactivator OCA-B. *Mol. Cell. Biol.* 15: 4115-4124.

Ma J. and Ptashne M. 1987a. Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* 48: 847-853.

Ma J. and Ptashne M. 1987b. A new class of yeast transcriptional activators. *Cell* 51: 113-119.

Mangelsdorf D. J. and Evans R. M. 1995. The RXR heterodimers and orphan receptors. *Cell* 83: 841-850.

Manning E. T., Ikehara T., Ito T., Kadonaga J. T., and Kraus W. L. 2001. p300 forms a stable, template-committed complex with chromatin: Role for the bromodomain. *Mol. Cell. Biol.* 21: 3876-3887.

Mapp A. K. and Ansari A. Z. 2007. A TAD further: Exogenous control of gene activation. *ACS Chem. Biol.* 2: 62-75.

Margolin J. F., Friedman J. R., Meyer W. K., Vissing H., Thiesen H. J., and Rauscher F. J., 3rd. 1994. Krüppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci.* 91: 4509-4513.

Marmorstein R., Carey M., Ptashne M., and Harrison S. C. 1992. DNA recognition by GAL4: Structure of a protein-DNA complex. *Nature* 356: 408-414.

Massari M. E. and Murre C. 2000. Helix-loop-helix proteins: Regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* 20: 429-440.

McStay B., Frazier M. W., and Reeder R. H. 1991. xUBF contains a novel dimerization domain essential for RNA polymerase I transcription. *Genes Dev.* 5: 1957-1968.

Merickel S. K., Sanders E. R., Vazquez-Ibar J. L., and Johnson R. C. 2002. Subunit exchange and the role of dimer flexibility in DNA binding by the Fis protein. *Biochemistry* 41: 5788-5798.

Miller J., McLachlan A. D., and Kiug A. 1985. Repetitive zinc-binding domains in the protein transcription factor lIIA from Xenopus oocytes. *EMBO J.* 4: 1609-1614.

Montminy M., Brindle P., Arias J., Ferreri K., and Armstrong R. 1996. Regulation of somatostatin gene transcription by cyclic adenosine monophosphate. *Metabolism* 45: 4-7.

Mujtaba S., Zeng L., and Zhou M. M. 2007. Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* 26: 5521-5527.

Naar A. M., Beaurang P. A., Zhou S., Abraham S., Solomon W., and Tjian R. 1999. Composite coactivator ARC mediates chromatindirected transcriptional activation. *Nature* 398: 828-832.

Nedialkov Y. A., Shooltz D. D., and Triezenberg S. J. 2003. Purification and protein interaction assays of the VP 16C transcription activation domain. *Methods Enzymol*. 370: 522-535.

Nikolov D. B. and Burley S. K. 1997. RNA polymerase Ⅱ transcription initiation: A structural view. *Proc. Natl. Acad. Sci.* 94: 15-22.

Nikolov D. B., Chen H., Halay E. D., Hoffman A., Roeder R. G., and Burley S. K. 1996. Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc. Natl. Acad. Sci.* 93: 4862-4867.

Norton J. D. 2000. ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis. *J. Cell Sci*. 113: 3897-3905.

Novatchkova M. and Eisenhaber F. 2004. Linking transcriptional mediators via the GACKIX domain super family. *Curr. Biol.* 14: R54-55.

Ogryzko V. V., Schiltz R. L., Russanova V., Howard B. H., and Nakatani Y. 1996. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* 87: 953-959.

Onate S. A., Tsai S. Y., Tsai M. J., and O' Malley B. W. 1995. Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science* 6: 1354-1357.

Pabo C. O., Sauer R. T., Sturtevant J. M., and Ptashne M. 1979. The repressor contains two domains. *Proc. Natl. Acad. Sci.* 76: 1608- 1612.

Parker D., Ferreri K., Nakajima T., LaMorte V. J., Evans R., Koerber S. C., Hoeger C., and Montminy M. R. 1996. Phosphorylation of CREB at Ser-133 induces complex formation with CREB-binding protein via a direct mechanism. *Mol. Cell. Biol.* 16: 694-703.

Pufall M. A. and Graves B. J. 2002. Autoinhibitory domains: Modular effectors of cellular regulation. *Annu. Rev. Cell Dev. Biol.* 18: 421-462.

Puig O., Caspary R, Rigaut G., Rutz B., Bouveret E., Bragado-Nilsson E., Wilm M., and Seraphin B. 2001. The tandem affinity purification (TAP) method: A general procedure of protein complex purification. *Methods* 24: 218-229.

Rachez C., Suldan Z., Ward J., Chang C. P., Burakov D., ErdjumentBromage H., Tempst P., and Freedman L. P. 1998. A novel protein complex that interacts with the vitamin D3 receptor in a liganddependent manner and enhances VDR transactivation in a cell- free system. *Genes Dev.* 12: 1787-1800.

Ramirez-Carrozzi V. and Kerppola T. 2001. Gel-based fluorescence resonance energy transfer (gelFRET) analysis of nucleoprotein complex architecture. *Methods* 25: 31-43.

Regier J. L., Shen F., and Triezenberg SI. 1993. Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator. *Proc. Natl. Acad. Sci.* 90: 883-887.

Rellahan B. L., Jensen J. P., Howcroft T. K., Singer D. S., Bonvini E., and Weissman A. M. 1998. Elf-1 regulates basal expression from the T cell antigen receptor ζ-chain gene promoter. *J. Immunol.* 160: 2794-2801.

Ren Y., Behre E., Ren Z., Zhang J., Wang Q., and Fondell J. D. 2000. Specific structural motifs determine TRAP220 interactions with nuclear hormone receptors. *Mol. Cell. Biol.* 20: 5433-5446.

Rigaut G., Shevchenko A., Rutz B., Wilm M., Mann M., and Seraphin B. 1999. A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 17: 1030-1032.

Rodriguez P., Braun H., Kolodziej K. E., de Boer E., Campbell J., Bonte E., Grosveld F., Philipsen S., and Strouboulis J. 2006. Isolation of transcription factor complexes by in vivo biotinylation tagging and direct binding to streptavidin beads. *Methods Mol. Biol.* 338: 305-323.

Rogatsky I., Wang J. C., Derynck M. K., Nonaka D. F., Khodabakhsh D. B., Haqq C. M., Darimont B. D.,

Garabedian M. J., and Yamamoto K. R. 2003. Target-specific utilization of transcriptional regulatory surfaces by the glucocorticoid receptor. *Proc. Natl. Acad. Sci.* 100: 13845-13850.

Rosenberg A. H., Lade B. N., Chui D. S., Lin S. W., Dunn J. J., and Studier F. W. 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* 56: 125-135.

Roth J. F., Shikama N., Henzen C., Desbaillets I., Lutz W., Marino S., Wittwer J., Schorle H., Gassmann M., and Eckner R. 2003. Differential role of p300 and CBP acetyltransferase during myogenesis: p300 acts upstream of MyoD and Myf5. *EMBO J.* 22: 5186-5196.

Sadowski I., Ma J., Triezenberg S., and Ptashne M. 1988. GAL4-VP16 is an unusually potent transcriptional activator. *Nature* 335: 563-564.

Saha S., Brickman J. M., Lehming N., and Ptashne M. 1993. New eukaryotic transcriptional repressors. *Nature* 363: 648-652.

Sambrook J. and Russell D. W. 2001. Mutagenesis. In *Molecular cloning: A laboratory manual*, 3rd ed., pp. 13.1-13.91. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sanjabi S., Hoffmann A., Liou H. C., Baltimore D., and Smale S. T. 2000. Selective requirement for c-Rel during IL-12 P40 gene induction in macrophages. *Proc. Natl. Acad. Sci.* 97: 12705-12710.

Sanjabi S., Williams K. J., Saccani S., Zhou L., Hoffmann A., Ghosh G., Gerondakis S., Natoli G., and Smale ST. 2005. A c-Rel subdomain responsible for enhanced DNA-binding affinity and selective gene activation. *Genes Dev.* 19: 2138-2151.

Schena M., Freedman L. P., and Yamamoto K. R. 1989. Mutations in the glucocorticoid receptor zinc finger region that distinguish interdigitated DNA binding and transcriptional enhancement activities. *Genes Dev.* 3: 1590-1601.

Shen F., Triezenberg S. J., Hensley P., Porter D., and Knutson J. R. 1996a. Critical amino acids in the transcriptional activation domain of the herpesvirus protein VP16 are solvent-exposed in highly mobile protein segments. An intrinsic fluorescence study. *J. Biol. Chem.* 271: 4819-4826.

Shen F., Triezenberg S. J., Hensley P., Porter D., and Knutson J. R. 1996b. Transcriptional activation domain of the herpesvirus protein VP16 becomes conformationally constrained upon interaction with basal transcription factors. *J. Biol. Chem.* 271: 4827-4837.

Sheppard H. M., Harries I. C., Hussain S., Bevan C., and Heery D. M. 2001. Analysis of the steroid receptor coactivator 1 (SRC1) -CREB binding protein interaction interface and its importance for the function of SRC1. *Mol. Cell. Biol.* 21: 39-50.

Shibata H., Spencer T. E., Onate S. A., Jenster G., Tsai S. Y., Tsai MI., and O' Malley B. W. 1997. Role of co-activators and co-repressors in the mechanism of steroid/thyroid receptor action. *Recent Prog. Horm. Res.* 52: 141-164; discussion 164-145.

Shikama N., Lutz W., Kretzschmar R., Sauter N., Roth J. F., Marino S., Wittwer I., Scheidweiler A., and Eckner R. 2003. Essential function of p300 acetyltransferase activity in heart, lung and small intestine formation. *EMBO J.* 22: 5175-5185.

Singh S. V., Leal T., Ansari G. A., and Awasthi Y. C. 1987. Purification and characterization of glutathione S-transferases of human kidney. *Biochem. J.* 46: 179-186.

Smith D. B. and Johnson K. S. 1988. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene* 67: 31-40.

Smith M. J., Charron-Prochownik D. C., and Prochownik E. V. 1990. The leucine zipper of c-Myc is required for full inhibition of erythroleukemia differentiation. *Mol. Cell. Biol.* 10: 5333-5339.

Sopta M., Carthew R. W., and Greenblatt J. 1985. Isolation of three proteins that bind to mammalian RNA polymerase II. *J. Biol. Chem.* 260: 10353-10360.

Strubin M. and Struhl K. 1992. Yeast and human TFII D with altered DNA-binding specificity for TATA elements. *Cell* 68: 721-730.

Studier E. W. and Moffatt B. A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189: 113-130.

Studier F. W., Rosenberg A. H., Dunn J. J., and Dubendorif J. W. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* 185: 60-89.

Tan S., Hunziker Y., Sargent D. F., and Richmond T. J. 1996. Crystal structure of a yeast TF II A/TBP/DNA complex. *Nature* 381: 127-151.

Tanaka M., Clouston W. M., and Herr W. 1994. The Oct-2 glutaminerich and proline-rich activation domains can synergize with each other or duplicates of themselves to activate transcription. *Mol. Cell. Biol.* 14: 6046-6055.

Tang H., Sun X., Reinberg D., and Ebright R. H. 1996. Protein-protein interactions in eukaryotic transcription initiation: Structure of the preinitiation complex. *Proc. Natl. Acad. Sci.* 93: 1119-1124.

Tansey W. P. and Herr W. 1995. The ability to associate with activation domains in vitro is not required for the TATA box-binding protein to support activated transcription in vivo. *Proc. Natl. Acad. Sci.* 93: 10550-10554.

Tansey W. P. and Herr W. 1997. Selective use of TBP and TF II B revealed by a TATA-TBP-TF II B array with altered specificity. *Science* 275: 829-831.

Tiu W. U., Davern K. M., Wright M. D., Board P. G., and Mitchell G. F. 1988. Molecular and serological characteristics of the glutathione S-transferases of Schistosoma japonicum and Schistosoma mansoni. *Parasite Immunol.* 10: 693-706.

Traven A., Jelicic B., and Sopta M. 2006. Yeast Gal4: A transcriptional paradigm revisited. *EMBO Rep.* 7: 496-499.

Triezenberg S. J. 1995. Structure and function of transcriptional activation domains. *Curr. Opin. Genet. Dev.* 5: 190-196.

Triezenberg S. J., Kingsbury R. C., and McKnight S. L. 1988. Functional dissection of VPl6, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev.* 2: 718-729.

Uesugi M., Nyanguile O., Lu H., Levine A. J., and Verdine G. L. 1997. Induced a helix in the VP16 activation domain upon binding to a human TAF. *Science* 277: 1310-1313.

Urrutia R. 2003. KRAB-containing zinc-finger repressor proteins. *Genome Biol.* 4: 231.

van Steensel B. and Henikoff S. 2000. Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat. Biotechnol.* 18: 424-428.

van Tilborg M. A., Lefstin J. A., Kruiskamp M., Teuben J., Boelens R., Yamamoto K. R., and Kaptein R. 2000. Mutations in the glucocorticoid receptor DNA-binding domain mimic an allosteric effect of DNA. *J. Mol. Biol.* 301: 947-958.

Vinson C., Acharya A., and Taparowsky E. J. 2006. Deciphering B-ZIP transcription factor interactions in vitro and in vivo. *Biochim. Biophys. Acta* 1759: 4-12.

Vo N. and Goodman R. H. 2001. CREB-binding protein and p300 in transcriptional regulation. *J. Biol. Chem.* 276: 13505-13508.

Wagner S. and Green M. R. 1993. HTLV-I Tax protein stimulation of DNA binding of bZIP proteins by enhancing dimerization. *Science* 262: 395-399.

Walker S., Greaves R., and O' Hare P. 1993. Transcriptional activation by the acidic domain of Vmw65 requires the integrity of the domain and involves additional determinants distinct from those necessary for TF II B binding. *Mol. Cell. Biol.* 13: 5233-5244.

Wang L., Xie J., and Schultz P. G. 2006a. Expanding the genetic code. *Annu. Rev. Biophys. BiomoL Struct.* 35: 225-249.

Wang J., Rao S., Chu J., Shen X., Levasseur D. N., Theunissen T. W., and Orkin S. H. 2006b. A protein interaction network for pluripotency of embryonic stem cells. *Nature* 444: 364-368.

Weiner M. P., Costa G. L., Schoettlin W., Cline J., Mathur E., and Bauer J. C. 1994. Site-directed mutagene-

sis of double-stranded DNA by the polymerase chain reaction. *Gene* 151：119-123.
Weiss M. J.，Yu C.，and Orkin S. H. 1997. Erythroid-cell-specific properties of transcription factor GATA-I revealed by phenotypic rescue of a gene-targeted cell line. *Mol. Cell. Biol.* 17：1642-1651.
Yang F.，Vought B. W.，Satterlee J. S.，Walker A. K.，Jim Sun Z. Y.，Watts J. L.，DeBeaumont R.，Saito R. M.，Hyberts S. G.，Yang S.，et al. 2006. An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. *Nature* 442：700-704.
Yuan C. X.，Ito M.，Fondell J. D.，Fu Z. Y.，and Roeder R. G. 1998. The TRAP220 component of a thyroid hormone receptor-associated protein (TRAP) coactivator complex interacts directly with nuclear receptors in a ligand-dependent fashion. *Proc. Natl. Acad. Sci.* 95：7939-7944.
Zhang X.，Krutchinsky A.，Fukuda A.，Chen W.，Yamamura S.，Chait B. T.，and Roeder R. G. 2005. MED1/TRAP220 exists predominantly in a TRAP/Mediator subpopulation enriched in RNA polymerase II and is required for ER-mediated transcription. *Mol. Cell* 19：89-100.

网络资源

http://elm.eu.org/：通过鉴定常规模式和使用基于环境的规则和逻辑过滤，采用真核线性模体（eukaryotic linear motif，ELM）资源预测真核蛋白中的功能位点。

http://harvester.fzk.de/：生物信息学位点搜索工具及资源。

http://tools.invitrogen.com/search/index.cfm?fuseaction=search.simplesearch&searchterm=baculovirus：来自 Invitrogen 的关于杆状病毒表达系统的技术资源和信息。

http://www.bdbiosciences.com/features/productsldisplay_product.php?keyID=72：来自 BD Biosciences 的关于杆状病毒表达系统的技术资源和信息。

http://www.clontech.comlproducts/detail.asp?tabno=2&catalog_id=631402&page=all：来自 Clontech 的关于杆状病毒表达系统的技术资源和信息。

http://www.clontech.com/products/technologies.asp?product_famiiy_id=1419&pagesource=home：来自 Clontech 的哺乳动物表达系统的技术资源和信息。

http://www.emdbiosciences.com/html/NVG/insectdirectindex.html：昆虫指导系统（Insect Direct system）提供一种在昆虫细胞中表达异源蛋白的无病毒方法。

http://www.emdbiosciences.com/html/NVG/pETTable.html：来自 EMD Biosciences 的基于 T7 启动子驱动系统的 pET 载体的技术资源和信息。

http://www.expasy.org/：ExPASy（专家蛋白分析系统）含有用于蛋白质结构域和结构分析的多种在线工具的链接。“蛋白组学和序列分析工具（proteomics and sequence analysis tools）”部分含有用于分析保守结构域和较小模体［即“模式和图谱搜索（pattern and profile searches）”］的链接。

http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Protein-Expression-and-Analysis/Protein-Expression/Mammalian-Expression.htnsl：来自 Invitrogen 的哺乳动物表达系统的技术资源和信息。

http://www.ncbi.nlm.nih.gov/：NCBI（国家生物信息中心）维护一个广泛使用的，并且基础广泛的数据库系列和搜索引擎。

http://www.neb.com/nebecomm/products/productE3000.asp：来自 New England Biolabs 的哺乳动物表达系统的技术资源和信息。

http://www.promega.com/vectors/mammalian_express_vectors.htm：来自Promega的哺乳动物表达系统的技术资源和信息。

http://www.sbg.bio.ic.ac.uk/—phyre/：该程序由Sternberg实验室开发，用于基于相关蛋白质的已知结构来预测基本结构的研究。通过最小化能量，该程序产生避免空间冲突的模型；空间冲突可能来自已知结构的蛋白质与同源蛋白质间的氨基酸差异。

http://www.stjude.org/brindle：该网站是Paul Brindle实验室建立的网站，该网站维护一个关于p300和CBP互作组的数据库。

http://www.ysbl.york.ac.ukl—emsley/coot/：面向晶体学对象的工具包（Crystallographic Object-Oriented Toolkit，Coot），允许蛋白质结构的复杂分析，包括蛋白质数据库文件的操作。

http:wwwl.gelifesciences.com/aptrix/upp01077.nsf/Content/Products?OpenDocument&parentid=976038&moduleid=38859&zone=Proteomics：来自GE Healthcare的用于GST融合的pGEX载体相关技术资源和信息。

11 DNA 的调控性转录因子结合

要点：

- 转录因子使用多种不同的机制识别它们的 DNA 结合位点。
- 许多真核生物的蛋白质协同结合 DNA 形成复合物或增强体。
- 可以利用几种简单的方法测量和定量 DNA 结合。
- 建立 DNA-蛋白质相互作用的模型对于了解转录因子与 DNA 的结合机制是必需的。

引言和概述

启动子的机制分析一般涉及确定序列特异性转录调控蛋白如何识别并结合它们的 DNA 位点的实验，不管这些蛋白质是单独地还是组合地结合 DNA。第 7 章讨论了 DNA 结合蛋白的鉴定，第 8 章讨论了确定 DNA-蛋白质相互作用生理相关性的标准。这里我们将讨论 DNA 识别的原理、如何鉴定 DNA 结合蛋白的高亲和识别位点，以及如何利用化合物和核酸酶探针研究 DNA-蛋白质相互作用。然后我们将详尽阐述研究多激活因子复合物或增强体形成的一些简单原理和方法。

目前的模型是：真核 DNA 结合蛋白通过不停地碰撞核 DNA 或在核 DNA 上滑行，直至遇到启动子内的功能性位点时，才与其生理位点结合（综述见 Jen-Jacobsen 1997；

Halford and Marko 2004)。功能性位点是被调控启动子上下文中介导转录因子生理作用的位点。而非功能性位点则由随机的非特异性序列或处于错误的上下文中的特异性识别序列构成。

细胞已经设计出三种策略来加速这种搜索，并赋予这种搜索特异性。首先，细胞中许多未转录的DNA被包装进染色质，这些DNA在很大程度上对调控因子来说是不可及的；其次，细胞核内DNA结合蛋白的浓度很高，足以克服来自非特异性DNA的显著竞争；最后，仅在合适的启动子环境中发生的蛋白质-蛋白质及DNA-蛋白质相互作用才会产生大量的结合能。

环境依赖的DNA-蛋白质及蛋白质-蛋白质相互作用对于位点的定位极为重要。以下两点必须要考虑。第一，在裸基因组DNA中，DNA结合蛋白能识别的位点的实际数目，远远超过生理性位点的个数。一个典型的6 bp DNA序列在人基因组出现约732 422次（换而言之，也就是3亿个碱基对除以4^6，6 bp位点可能组合的个数）。因为基因组中只有约25 000个蛋白质编码基因（http://www.ornl.gov/sci/techresources/Human_Genome/faq/genenumber.shtml），并且不可能任何调控因子（除聚合酶II[Pol II]通用机器）都与所有这些位点结合，对于特定的因子来说，识别位点的实际数目很可能超过生理相关位点或功能性位点的数目。第二，转录因子通常属于在体外识别相关或相同位点的蛋白质家族（Wolberger 1999；Luscombe et al. 2000）（例如，Ets家族成员；见Hollenhorst et al. 2004）。事实上，真核基因组通常将其7%的编码序列奉献给DNA结合蛋白。这样，人类基因组编码大约1800个因子和很多多基因家族（同源结构域、Ets、锌指结构及亮氨酸拉链）。由于转录因子家族成员在不同启动子上的精心策划的行动对真核细胞的恰当运作和发育是关键的，所以这提出一个问题，即生理特异性怎样被传递给DNA-蛋白质相互作用。

曾经援引增强体理论来解释蛋白质如何能够取得恰当的特异性（见第1章）（Echols 1986；Grosschedl 1995；Carey 1998；Merika and Thanos 2001；Panne 2008）。这一概念是，启动子/增强子中各位点的排布及结合这些位点的调控蛋白的特定集合形成了一个独特的蛋白质-DNA及蛋白质-蛋白质相互作用网络。最终结构的能量或稳定性取决于结合位点的精确安置和合适的调控因子与这些位点的结合。最终目标是以最低的自由能组装成一个稳定的复合物，非常像用智力玩具的组成零件组装成一个智力玩具。

从结构生物学的研究中产生了两个概念，用以解释增强体是如何形成的。一个概念显示蛋白质-蛋白质相互作用增加蛋白质组合物对相邻识别位点的亲和力，如图11.1A所示的白细胞介素-2（IL-2）启动子上活化T细胞核因子（NFAT）、Fos和Jun的晶体结构（Chen et al. 1998）。另一个概念显示通过一个调控因子的结合产生的DNA构象有利于另一个调控因子的结合，如图11.1B所示的β-干扰素（IFN-β）增强子上激活转录因子2（ATF-2）/Jun和干扰素调控因子-3（IRF-3）的晶体结构（Panne et al. 2004）。在这两种结合方式中，对蛋白质的空间约束和DNA结构的改变都有利于独特构造的形成，这种独特构造也对特异性作出贡献。事实上，包含ATF-2/c-Jun、IRF-3/IRF-7和核因子-κB（NF-κB）的整个IFN-β增强体，揭示了一个连续的蛋白质表面，该表面被设计用于识别构成增强子识别序列的复合DNA结合元件（Panne et al. 2007，2008）。

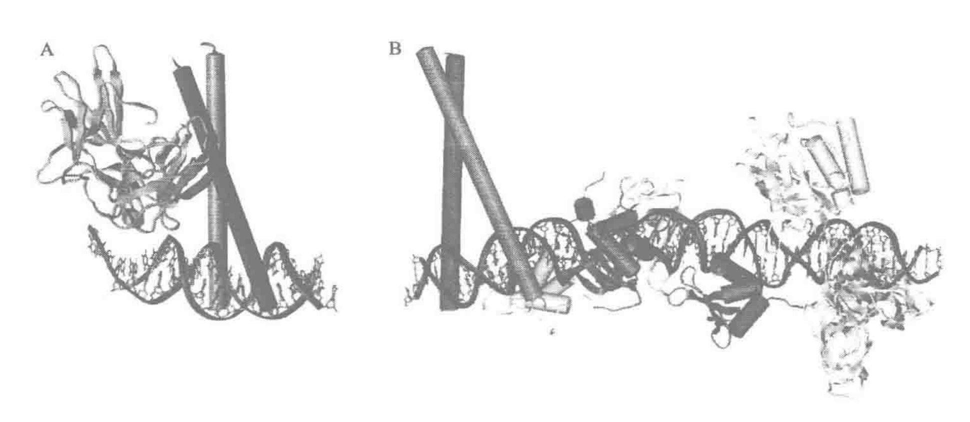


图 11.1 典型的增强体的结构。(A) Fos-Jun-NFAT。一个实例，其中蛋白质-蛋白质相互作用使蛋白质协同结合到邻近 DNA 位点上（经 MacMillan Publishers Ltd 允许，改编自 Chen et al. 1998）。(B) ATF-2/Jun、IRF-3、IRF-7 和 NF-κB 结合在 IFN-β 增强体上（经允许，改编自 Panne et al. 2007，© Elsevier）。DNA 构象的改变有利于 ATF-2/Jun 与两分子的 IRF-3 和 IRF-7 及 NF-κB 协同结合。

该增强体模式阐明了激活因子是如何区分其生理性位点和非生理性位点的。首先，通过与复合物中其他蛋白质的协作，激活因子对其生理性位点具有更高的亲和力。据推测，在标准条件下，由于激活因子的浓度、蛋白质-蛋白质相互作用及 DNA 变形之间正确的平衡，只有蛋白质的特定组合物才能结合并组装增强体。其次，这种模式解决了相关序列偏好的悖论。尽管有些调控蛋白可能识别同一个序列，但随后的定向相互作用和复合物的最终自由能“选择”合适的因子。选择性的另一个层面是，原则上，增强体本身可生成一个表面，该表面与位于聚合酶 II 通用机器的不同组分上的一个表面互补。只有形成适当的界面时，增强体才会使 DNA 环状突出并将聚合酶 II、共激活因子及通用因子募集到启动子上。事实上，通过使用这样的机制，通用聚合酶 II 转录机器才可以凭借相互协作的相互作用协助组装增强体。

根据这一假说，协作性的启动子环境依赖性相互作用是区分非生理性位点和功能性位点的驱动力。使 DNA 变形或弯曲的建筑蛋白（architectural protein），如 A 或 B 家族的高迁移组（HMG）蛋白（Agresti and Bianchi 2003），可能是这种复合物的必要组成部分。虽然这些蛋白质不会以与激活因子相同程度的特异性结合（在许多情况下，它们被视为非特异性的），但它们可能允许某种组合相互作用，只不过这种相互作用受到发生于其间的 DNA 在其构象保持长度（persistence length）内有限的柔韧性的制约，以及结合的蛋白质的大小和柔韧性的限制。

应当指出，只有少数增强体的例子已在体外和体内研究中被证实。增强体的高稳定性与一些体内 DNA 结合蛋白（如糖皮质激素受体和 NF-κB）的瞬时的和动态的结合形成鲜明的对比（McNally et al. 2000；Bosisio et al. 2006；Voss and Hager 2008）。因此，随着对越来越多的体外和体内研究的比较，将在体外观察到的增强子结合蛋白的协作性高亲和力结合与在活细胞内观察到的相对瞬时结合协调一致将是非常重要的。

本章考察用于理解启动子识别的基本环节的生化方法及策略。我们首先聚焦于单克隆调控蛋白是如何识别其结合位点的，随后总结了用于研究DNA弯曲及协作结合的方法，DNA弯曲和协作结合是产生增强体复合物所必需的两个现象。

实验策略

DNA-蛋白质相互作用的一般理论和实例

DNA识别的理论

调控蛋白对DNA位点的识别受到它与碱基的特异性相互作用及与糖/磷酸骨架的非特异性相互作用的影响。虽然围绕使用这一简化术语描述DNA结合有一些轻微的争议，但它提供了一个可在调控环境的基础上进行细化的框架。在一个典型的相互作用中，调控蛋白（P）和DNA位点（S）与蛋白质位点（PS）复合物之间处于平衡。该平衡可表示为PS⟷P+S，其平衡解离常数$K_d = k_{off}/k_{on} =$［P］［S］/［PS］，其中，［P］、［S］和［PS］是平衡时各种组分的浓度。这一平衡常数，也定义为速率常数的比值，考虑到所有的促进结合的焓能量和熵能量，包括定位位点的能量损耗（即$\Delta G=\Delta H-T\Delta S$）。还可以根据自由能定义该$K_d$，使用吉布斯自由能方程：$\Delta G = RT \ln K_d$（至于该方程如何对生物反应产生影响，见Dill 1997）。请注意此处我们使用的是一种变种的吉布斯自由能方程。大多数教科书中列出的是相对于结合或平衡常数（$\Delta G = -RT \ln K_a$）的自由能方程。因为我们这里所指的能量是从解离常数计算而来的，故正确的术语是$\Delta G = -RT \ln (1/K_d)$，该公式等同于$\Delta G = RT \ln K_d$。

单个蛋白质-DNA相互作用的能量和特异性产生于氨基酸侧链的独特立体定向排列，这些排列与处于DNA大沟或小沟中的碱基所展示的化学基团的排列在化学上和空间上形成互补。每种化学相互作用提供一个量子的自由能；根据前面吉布斯公式所描述的自由能（ΔG）和K_d之间的对数关系，与理想位点的每点偏离都会引起位点亲和力的剧烈下降，这种对数关系描述了一种增强特异性的机制。上述互补形式称为直接读出。

DNA识别的另一种形式称为间接读出，这种形式涉及一种蛋白质基于DNA的二级结构、构象或变形结合特定序列的能力。如果识别位点偏离最适位点，则在该位点上DNA的固有变形就会受到影响。产生的K_d值改变可能远大于通过从简单的化学相互作用丧失的能量预测的K_d改变。例如，*Eco*RI的GAATTC识别位点依赖于特定序列的排列，以适应一个非常复杂的与催化密切相关的变形，该位点中单个碱基的替换就能改变结合的ΔG，并随后提高K_d值（Lesser et al. 1990；Jen-Jacobsen 1997）。

除了氨基酸侧链与碱基对的暴露基团的特异性相互作用外，与B-DNA螺旋相对均匀的糖/磷酸骨架的非特异性氢键及范德华力相互作用也对结合能有贡献。晶体学研究表明，在识别复合物中，这些相互作用通常提供大量的自由能，尽管不产生特异性（Luscombe et al. 2000）。这些非特异性相互作用也可能通过允许该蛋白质结合并沿DNA骨架的等电表面滑行，同时持续地尝试大沟和小沟，以找到允许蛋白质稳定停泊的特定化学识别标志，从而促进蛋白质初步定位于一个特定的位点。除了上述的焓效应

外，来源于蛋白质结合时从位点上有序的水和盐离子释放的熵贡献，据认为也有助于蛋白质-DNA 的相互作用（Ha et al. 1989；Jen-Jacobsen 1997）。

相互作用的化学基础

现在，我们更加仔细检查一下对 DNA 结合的主要焓贡献——氨基酸侧链，偶尔也有骨架上的酰胺和羰基、与糖/磷酸骨架和暴露在大沟或小沟中的碱基对的化学基团的相互作用（Luscombe et al. 2001）。下面列出了三类在大沟和小沟中的相互作用。

1. 与沿 DNA 螺旋的碱基对和磷酸的暴露边缘形成的氢键（图 11.2B 和图 11.2C）。由各种各样的氨基酸侧链（图 11.2B）与肽骨架中的酰胺基和羰基（图 11.2C）所支持的氨基酸-DNA 氢键相互作用。所有蛋白质-DNA 复合物都使用此策略。蛋白质-DNA 复合物中约 1/6 的相互作用是直接氢键。

2. 水分子可充当一种氨基酸和碱基对之间的氢键桥梁［或色氨酸操纵子阻遏物（Trp-R）共复合物；Otwinowski et al. 1988)］（图 11.2D）。在某些情况下，（如 Trp-R），这些接触能够给相互作用增添特异性，而在其他情况下，水分子仅仅是氨基酸和 DNA 之间的简单桥梁。蛋白质- DNA 复合物中约 1/6 的相互作用是水介导的氢键。

3. 范德华力相互作用（图 11.2A）。值得注意的是，蛋白质-DNA 复合物中近 2/3 的相互作用是范德华力相互作用。许多氨基酸的堆积相互作用沿 DNA 骨架发生。其中，范德华力相互作用就是例子，在这种范德华力相互作用中氨基酸侧链插入到两个碱基之间。这种插入事件主要是在 DNA 弯曲蛋白中观察到的，典型的例子是 TATA 结合蛋白（TBP)，该蛋白质将两个苯丙氨酸插入到 TATA 盒的第一个和最后一个碱基对之间，这种插入会产生 80° 的弯曲，从而解开双螺旋并拓宽小沟。HMG 蛋白质也使用类似的策略（Werner and Burley 1997)。

α 螺旋在 DNA 识别中的作用

一种蛋白质如何特异性地识别 DNA 呢？α 螺旋的大小和形状（圆柱体形，主链直径 5Å，含丙氨酸侧链时主链直径约 8.5Å）非常适合 B 型 DNA 的大沟（12Å 宽，6Å 深)，并且毫无意外，大多数 DNA 结合蛋白都采用这种策略（图 11.2E)。因此，氨基酸侧链的化学多样性和柔韧性，以及螺旋轴的旋转赋予 α 螺旋大量潜在的识别表面，用于结合特异性 DNA 序列。晶体研究也表明，这种所谓的识别 α 螺旋在相对于 DNA 骨架轴心的大沟中的安置，在不同的调控蛋白中差别很大。但是，由于几乎没有证据能说明一个孤立的 α 螺旋能够独立识别（Pabo and Sauer 1992)，因此，已经提出由其他蛋白质元件介导的、与糖/磷酸骨架的相互作用，对于螺旋在其位点上的正确定位是必需的。因此，一旦处在了合适的位点上，识别螺旋就被蛋白质支架（包括邻近的 α 螺旋和 β 折叠片）及与 DNA 骨架非特异性相互作用的复杂网络所稳定。虽然对使用大沟中 α 架螺旋的结构有所偏好，但在 DNA 识别中也有许多使用 β 折叠片或蛋白质环的例子（Luscombe et al. 2000)。而且，许多蛋白质除了大沟相互作用外，还利用柔韧的肽链伸展到达 DNA 的背面并与小沟接触，从而增强识别的特异性。

Thornton 及其同事仔细分析了 240 个高分辨率的蛋白质-DNA 复合物，生成一个 DNA 识别结构域的分类系统和分类表。8 类 DNA 识别结构域分别是：螺旋-转角-螺旋（HTH)、锌协调（zinc-coordinating)、拉链型（zipper type)、其他 α 螺旋、β 折叠片、

图 11.2 (A～D). DNA结合相互作用的化学基础。E. 大沟中的一个 α 螺旋［A 经允许改编自 Kissinger et al. 1990，© Cell Press。B 经允许改编自 Branden and Tooze 1991，© Garland Science/Taylor & Francis，LLC。C 经允许改编自 Jordan and Pabo 1988，© 1988 美国科学促进会 (AAAS)。D 经允许改编自 Feng et al. 1994，© 1994 美国科学促进会 (AAAS)。E 经允许改编自 Ptashne 1992，© Blackwell Science］。

β发夹带、其他及酶。这些类别的每一种基于其二级结构比对程序进一步细分为不同的同源家族。例如，至少有4种不同的锌协调家族。Luscombe 等（2000）报道了对这些家族的描述及其结合的方式。对于那些正在进行 DNA-蛋白质相互作用建模的或希望查

看不同家族三维结构的人，研究者们已经生成了一个非常优秀的网站（http://www.biochem.ucl.ac.uk/bsm/prot_dna/prot_dna_cover.html）。

大沟和小沟的特异性

尽管上面我们已经描述了 DNA 位点识别的能量学，但是没有讨论 DNA 上的互补表面，该互补表面允许 DNA 以序列特异的方式与蛋白质表面对接。也就是说，不同的 DNA 序列如何自我标记，以便被调控蛋白识别？虽然在其大小和扭角上显示序列依赖性变化，但 DNA 的构象和形状在结构上相对一致，正如磷酸骨架及糖构象那样。相反，特异性通常是由来自碱基的化学基团向大沟和小沟的序列依赖性突入赋予的。

当 A 和 T 或 C 和 G 之间通过氢键键合形成连接 DNA 反向平行链的碱基对时，所有 4 个碱基都暴露出化学基团，这些化学基团要么是多化合价的，要么不参与配对。可以想象这些暴露的基团可能与调控蛋白表面的氨基酸侧链或肽骨架发生键合。然而，只有当大沟和小沟中可及的氢键供体及受体的排布随 DNA 序列发生显著变化时，这些化学基团才可能赋予特异性。Seeman 等（1976）通过考察不同碱基对中化学基团的排布对这一问题进行了仔细的考虑。他们推断每个碱基对在大沟中都展示了一个独特的化学基团的三维模式。多碱基识别位点显示一个组合的并且因此更为复杂的阵列，这样，一个阵列的独特性随位点大小的增大而增强，这一变化在大沟中比在小沟中更加明显。

阵列如何发生变化呢？碱基对之间有 4 种组合：AT、TA、GC 和 CG。每个碱基对除了含有以氢键键合从而形成碱基对的化学基团外，还展示某些未键合的化学基团；这些未键合的化学基团分为三类：①氢键受体（ac）；② 氢键供体（do）；③与胞嘧啶 C5 氢（vdw-h）的范德华接触，或更频繁地的与胸腺嘧啶 C5 上突出的甲基（vdw-me）的范德华接触（图 11.3）。在 AT 的情况中，大沟展示空间排列 ac/do/ac/vdw-me，而 TA 展示 vdw-me/ac/do/ac。同样，GC 碱基对展示 ac/ac/do/vdw-h，而 CG 碱基对展示 vdw-h/do/ac/ac。因此，关于化学基团的排列，每种碱基对都是独特的。相比之下，TA 碱基对和 AT 碱基对的小沟都展示 ac/ac 排列，同样地，GC 和 CG 都展示 ac/do/ac 排列。因此，在大沟中有显著的多样性，并且 AT 可区别于 TA，GC 可区别于 CG。然而，小沟相对裸露，并且 AT 和 TA 没有区别，GC 和 CG 也没有区别，只有 AT/TA 可区别于 GC/CG。由于这一原因及可及性问题，大沟就成为序列特异性的主要靶点。

虽然大多数 DNA 结合蛋白都偏好大沟，但也确实存在清楚描述的特异性小沟识别的实例。有些蛋白质将大沟和小沟识别作用相联系，而其他蛋白质主要与小沟相互作用。将大沟和小沟识别作用相结合的途径包括 Oct-1、POU 和 MATα-2 同源结构域等。这些蛋白质通过同源结构域 HTH 模体和大沟接触，通过伸展的氨基末端臂与小沟接触（如见 Wolberger et al. 1991；Klemm et al. 1994；Garvie and Wolberger 2001）。而有些蛋白质专门识别小沟，这种小沟特异性机制的一个很好的例子就是 TBP 和 TATA 框的结合（综述见 Patikoglou et al. 1999）。TBP 利用 10 个反向平行的 β 折叠片层形成一个凹陷的底面，该底面识别并结合小沟中的 TATA 位点。具有 HMG 结构域的蛋白质也利用与小沟的广泛接触。淋巴增强子结合因子 1（LEF-1）是一种 HMG B-box 蛋白，该蛋白质使 DNA 螺旋在小沟处扭曲和弯曲，从而促进 T 细胞受体 α（TCR-α）增强子上增强体的形成（Grosschedl et al. 1994；Love et al. 1995）。该 HMGB 结构域是一个 L

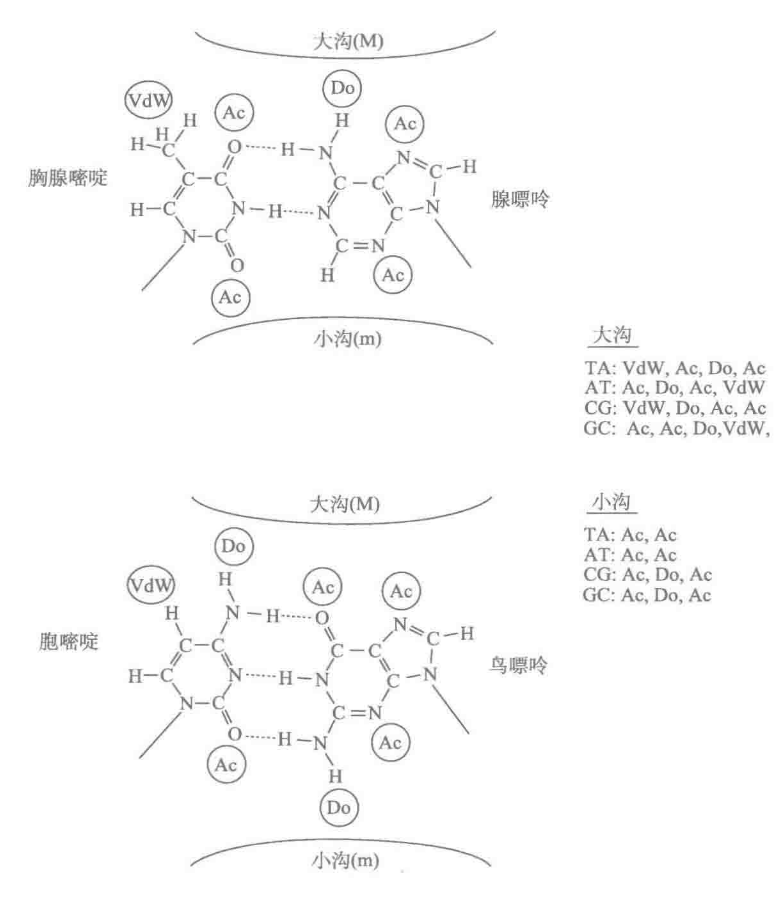


图 11.3 氢键供体和受体（经允许改编自 Watson et al. 2004，Benjamin Cummings Publishing，© 2004 Pearson Education，Inc.，Glenview，IL）。

形分子，可以塑造小沟，使小沟适应它的三个 α 螺旋（综述见 Thomas and Travers 2001）。TBP 和 HMGB 结构域都采用插入性氨基酸作为其特异性机制的构件。这些例子表明对于几个主要类型的 DNA 结合蛋白来说，小沟结合是它们序列识别的重要组成部分。

目前，蛋白质对 DNA 序列的识别过程尚没有通用密码。氨基酸侧链或主链酰胺基和羰基基团参与与各种各样的碱基或磷酸骨架的序列特异性相互作用。不管怎样，对来自蛋白质数据库（Protein DataBase）的 129 个结构中存在的化学键进行统计分析，的确揭示出了一些有趣的现象（Luscombe et al. 2001）。经过滤从数据库中除去结构相关复合物的数据之后，这些研究人员分析了 1111 个氢键、3576 个范德华接触和 821 个水介导键。表 11.1 显示了化学键的分布，包括每个氨基酸与每个碱基或糖/磷酸骨架之间形成的键，以及与这些键相对的仅仅基于随机对接在理论上可能的键。正如预想的一样，

表 11.1　根据参与的氨基酸和 DNA 碱基或骨架基团的氢键分布

氨基酸			DNA 碱基								DNA 骨架				合计	
			胸腺嘧啶(T)		胞嘧啶(C)		腺嘌呤(A)		鸟嘌呤(G)		糖		磷酸			
精氨酸	ARG	R	**24**	**(2.5)**	8	(2.0)	**19**	**(4.2)**	**98**	**(4.0)**	8	(1.9)	**218**	**(49.9)**	**375**	**(64.7)**
赖氨酸	LYS	K	**9**	**(4.4)**	6	(3.4)	**4**	**(7.4)**	**30**	**(7.1)**	3	(3.2)	**109**	**(86.7)**	**165**	**(112.6)**
丝氨酸	SER	S	3	(3.0)	2	(2.2)	1	(5.0)	12	(4.6)	2	(2.1)	**91**	**(57.3)**	**113**	**(74.4)**
苏氨酸	THR	T	5	(2.4)	3	(2.0)	4	(4.2)	—	(4.0)	1	(1.8)	**79**	**(49.2)**	**92**	**(63.9)**
天冬酰胺	ASN	N	7	(3.6)	10	(2.7)	**18**	**(6.0)**	7	(5.8)	3	(2.6)	**43**	**(70.7)**	88	(91.8)
谷氨酰胺	GLN	Q	2	(2.2)	2	(1.7)	**16**	**(3.8)**	6	(3.6)	2	(1.7)	**42**	**(44.8)**	70	(58.1)
甘氨酸	GLY	G	1	(3.2)	4	(2.4)	—	(5.4)	6	(5.1)	1	(2.4)	29	(63.3)	**41**	**(82.2)**
组氨酸	HIS	H	—	(0.8)	1	(0.6)	1	(1.5)	12	(1.4)	—	(0.7)	26	(18.3)	**40**	**(23.7)**
酪氨酸	TYR	Y	—	(1.2)	2	(1.0)	—	(2.1)	1	(2.0)	1	(1.0)	35	(25.7)	39	(33.4)
丙氨酸	ALA	A	1	(2.5)	1	(2.0)	—	(4.2)	1	(4.0)	—	(1.9)	**17**	**(49.8)**	**20**	(64.6)
谷氨酸	GLU	E	—	(3.8)	**10**	**(3.0)**	1	(6.5)	1	(6.2)	—	(2.9)	**6**	**(76.2)**	**18**	(99.0)
异亮氨酸	ILE	I	—	(0.7)	—	(0.5)	—	(1.3)	3	(1.2)	—	(0.6)	11	(15.9)	14	(20.7)
天冬氨酸	ASP	D	—	(4.5)	**5**	**(3.4)**	2	(7.5)	2	(7.2)	—	(3.3)	**2**	**(88.3)**	**11**	**(114.7)**
缬氨酸	VAL	V	—	(1.2)	—	(1.0)	—	(2.0)	—	(2.0)	—	(0.9)	**8**	**(24.5)**	**8**	**(31.8)**
半胱氨酸	CYS	C	—	(0.2)	1	(0.2)	—	(0.5)	—	(0.5)	—	(0.3)	4	(6.7)	5	(8.7)
苯丙氨酸	PHE	F	—	(0.6)	—	(0.5)	—	(1.1)	1	(1.1)	—	(0.5)	4	(14.4)	5	(18.6)
亮氨酸	LEU	L	—	(1.5)	—	(1.1)	—	(2.6)	—	(2.5)	—	(1.2)	**5**	**(30.8)**	**5**	**(40.0)**
甲硫氨酸	MET	M	1	(0.4)	—	(0.3)	—	(0.7)	—	(0.7)	—	(0.3)	3	(9.1)	4	(11.8)
色氨酸	TRP	W	—	(0.3)	—	(0.2)	—	(0.7)	—	(0.6)	—	(0.3)	3	(8.7)	3	(11.3)
脯氨酸	PRO	P	—	(3.5)	1	(2.7)	—	(6.0)	—	(5.7)	—	(2.6)	—	**(70.0)**	**1**	**(90.9)**
合计			53	(42.5)	**56**	**(33.0)**	66	(73.4)	**180**	**(69.4)**	21	(32.2)	**735**	**(860.3)**	1111	(1111)

注:经牛津大学出版社许可,转载自 Luscombe et al.,2001。

氨基酸显示在左栏中,碱基、糖和磷酸基团沿顶部显示。氨基酸和碱基按它们形成的相互作用的个数排序。来自随机蛋白质-DNA 对接的键的预期个数放在括号中,并且用 χ^2 检验评价观察数和预期数之间的离散程度。偏离 $P>0.9999$ 的预期分布的条目以粗体显示。由于 χ^2 检验需要一个大于 4 的预期值,某些条目根据有相似侧链的氨基酸被汇总:精氨酸和赖氨酸、天冬酰胺和谷氨酰胺,以及天冬氨酸和谷氨酸。其他无法合理组合的氨基酸没有被检验。

已知相互作用的2/3都是与骨架形成的，这些相互作用参与稳定蛋白质-DNA的相互作用，但并不产生特异性。在碱基接触中，鸟嘌呤参与大部分的氢键，这和预想的相同；鸟嘌呤有最多的暴露的氢键供体和受体（即相对于腺嘌呤有3个，胞嘧啶和胸腺嘧啶各有2个，鸟嘌呤却有4个）（图11.3）。正如预想的一样，精氨酸、赖氨酸、苏氨酸及丝氨酸，这些带正电的氨基酸和（或）极性氨基酸，都参与了与碱基的大部分键合，带负电荷的氨基酸和疏水性氨基酸却很少参与。在考察的355个氢键中，有131个是单氢键，有120个双氢键，也就是说，氨基酸上的两个化学基团与一个碱基或者一个单个的碱基对中的两个化学基团键合。这种二齿键增强了相互作用的特异性，因为与由氨基酸侧链和碱基边缘的化学几何结构造成的单键相比，这种二齿键存在的可能性更少。精氨酸-鸟嘌呤相互作用是最常见的。值得注意的是，这些相互作用中有121种被认为是复杂的，因为它们跨越同一DNA链上或不同DNA链上的两个或两个以上的碱基阶梯，并且主要使用精氨酸和赖氨酸。一个氨基酸跨越一个碱基阶梯并且还与邻近碱基相互作用的能力，让蛋白质得到一种基本的识别密码，因为碱基对组合代表了最基本的DNA序列。对于那些对碱基配对规则及这些碱基配对规则在DNA-蛋白质识别中可能如何产生特异性感兴趣的研究人员，可以参见下面的网址 http://www.biochem.ucl.ac.uk/bsm/sidechains/index.html.

单体和二聚体：能量学和调控考量

DNA识别单元的重复是自然界用以设计DNA结合蛋白最常见的策略之一（图11.4A）（见Garvie and Wolberger 2001）。采用这种重复的方式有两种：①二聚化或形成另一种更高级的结构，如三聚体或四聚体（图11.4AII）；② 在单一多肽链中DNA识别单元的多聚化（图11.4AIII）。因为亲和力与结合自由能呈指数相关（$\triangle G = RT \ln K_d$），通过加倍识别单元的个数来加倍结合能，这种结合能的加倍使二聚体相对于单体，或带有串联识别单元的单体相对于具有单一识别单元的单体的亲和力呈乘法增加。二聚体（如酵母GAL4和GCN4）分别与17 bp和7 bp位点结合，每个位点显示假二倍旋转对称（Carey et al. 1989；OUphant et al. 1989；Ellenberger et al. 1992）。晶体结构研究表明，在蛋白质与DNA结合过程中，每个单体识别旋转对称位点的一半。图11.5A显示了GCN4的结构，GAL4的结构在本章后面的部分中将会进行详细的介绍。

除了增加结合亲和力，二聚体还能将二聚化作为一种策略加强位点定位。原理上，两个单体在溶液中可二聚化，并且之后结合DNA，或者作为单体可顺序结合该位点并在DNA上二聚化。交联、荧光共振能量转移（FRET）及竞争性DNA对结合的差异影响的动力学分析表明，碱性亮氨酸拉链（bZIP）和HLH家族相关成员都更喜欢单体途径。两个单体的共价连接实际上减缓了蛋白质的位点定位能力。其基本原理是，虽然加倍接触个数提高了蛋白质的特异性亲和力，但它也增加了非特异性亲和力。尽管特异性亲和力与非特异性亲和力的比值有所增加，但二聚体从非特异性位点上解离的速度要比单体慢得多。借助作为单体的有序结合，降低非特异性亲和力，并提高“关闭”速率，由此加速锁定功能性生理位点的动力学（见Kohler and Schepartz 2001a，b及其中的参考文献）。

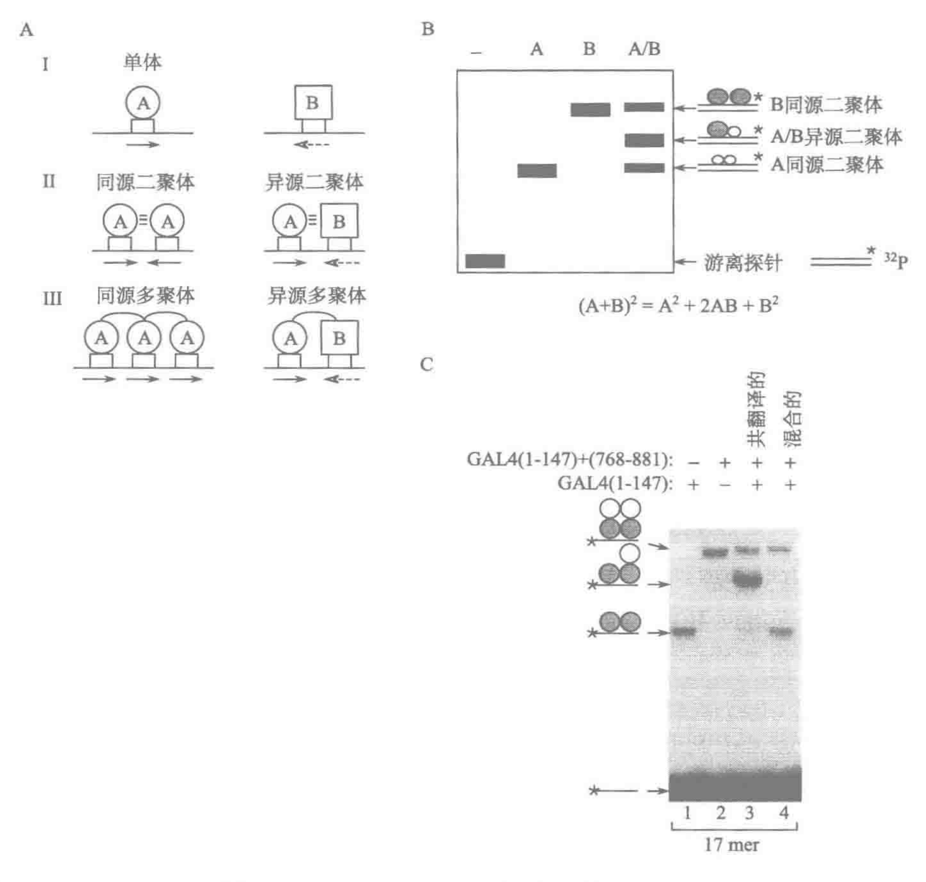


图 11.4　同源二聚体分析。(A) 重复 DNA 结合单元的不同策略。(B) 理论。(C) GAL4-衍生蛋白的实验。(C 经允许，修改自 Carey et al. 1989，© Academic Press Ltd.)

除了考虑到结合动力学因素外，许多蛋白质通过形成异源二聚体增加其调节的多样性，其中每个单体识别一个半位点。这一策略有双重目的。由于异源二聚体中每个单体都会有不同的序列偏好，从而使异源二聚体可识别含有非对称性半位点的位点。或者异源二聚体中的两个搭档可能有相同的 DNA 结合特异性，但具有不同的调控特性。在这方面有很好的例子，包括维甲酸 X 受体 (RXR) /视黄酸受体 (RAR) 异源二聚体与启动子上一个非对称直接重复的结合；RXR 识别 5′ 半位点，RAR (或另一个搭档) 识别 3′ 半位点 (图 11.5B)。依赖于 RXR 的搭档，异源二聚体可调控几种不同的直接重复的转录 (综述见 Mangelsdorf and Evans 1995)，并对不同组合的配体作出应答。Max 蛋白产物也能与多种相关蛋白质结合。Myc/Max 和 Mad/Max 异源二聚体与一个保守的 6 bp 调控位点结合，但是这两个异源二聚体具有非常不同的调控效应。Myc/Max 二聚体激活细胞周期依赖性基因，而 Mad/Max 二聚体则抑制这些相同的基因。另一个例子是真核 Jun 蛋白，它既能以同源二聚体 (Jun-Jun) 形式与 DNA 结合，又能以异源二聚体 (如 Jun-Fos、Jun-CREB、Jun-ATF-2) 形式与 DNA 结合。Jun 同源二聚体与 AP-1 启

动子元件微弱地结合，而 Jun-Fos 异源二聚体则与 AP-1 启动子元件紧密结合。Jun 和其他家族成员形成异源二聚体，这些二聚体与范围很大的一些列位点结合（Chinenov and Kerppola 2001）。

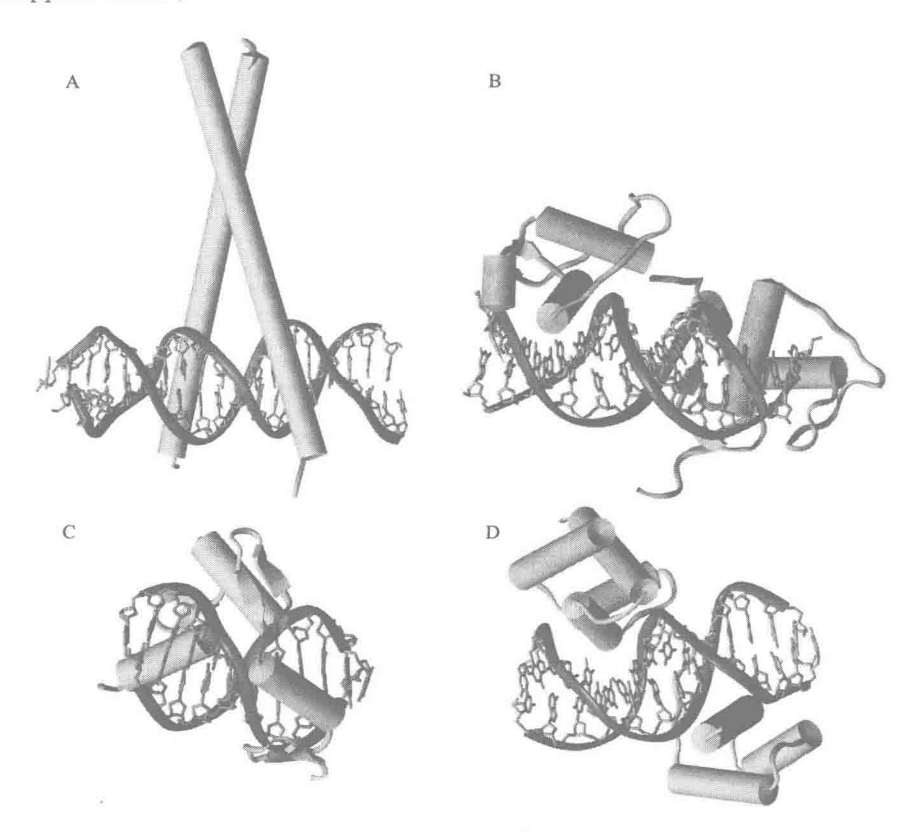


图 11.5　DNA 结合结构域的多聚化。（A）与其 DNA 位点结合的 GCN4 同源二聚体的结构描绘。（B）与其 DNA 位点结合的 RXR/RARA 异源二聚体的结构描绘。（C）与其 DNA 位点结合的 Zif268 的结构描绘。（D）与其 DNA 位点结合的 Oct-1 的结构描绘。B 和 C 中未标出锌原子（由 John Heiss [UCLA] 利用 Molecular Graphics structure modeling computer program Insight II 绘制）。

单一多肽中 DNA 结合模体的多聚化是另一种强大的用于增强特异性的方法。多聚化蛋白质可以含有多个相同的识别单元，也可以含有多个具有截然不同的模块化结构的单元。锌指蛋白（如 Zif268）就是一个很好的第一种情况的例子。典型的锌指结构中含有两个与一个 α 螺旋连接的反向平行的 β 折叠片，该螺旋中的两个半胱氨酸残基和两个组氨酸残基与第二个 β 折叠片像四面体一样协调一个锌离子。在 Zif268 中，三个 C-C-H-H 手指识别 DNA 双螺旋上相似的 3 bp 串联序列模体（图 11.5C）。锌指被精心设计，通过形成围绕 DNA 双螺旋的“C”形结构，恰好将 Zif268 的识别螺旋放入大沟中（Pavietich and Pabo 1991）。锌指结构需要识别不同的 DNA 序列，最近的研究工作表明，锌指结构可被重编程以识别截然不同的序列，从而生成结合独特序列的一系列独特的串联连接手指。

POU 结构域同源盒蛋白（如 Oct-1）采用结合结构域多聚化的替代方法（图 11.5D）。Oct-1 含有两个特异性 DNA 识别单元：POU 同源结构域和 POU 特异结构域。POU 同源结构域是一个 3-螺旋亚基，其中螺旋 2 和螺旋 3 形成 HTH 模体，并通过氨基末端臂与小沟结合。POU 特异结构域含 4 个 α 螺旋，这些螺旋使用的是与双螺旋广泛的碱基和磷酸骨架接触。这两个结构域一起与一个 8 bp 位点结合，并且两个结构域都是有效结合所必需的（Klemm et al. 1994）。

确定寡聚态的策略：确定一种蛋白质在溶液中或与其位点相结合时的寡聚态，是一项相对简单的任务。下面列出了 5 种通用的研究方法，根据需要解决的具体问题，这些方法有简单的，也有比较复杂的。

1. *化学及核酸酶足迹技术*：这些技术揭示一种二聚化蛋白是否对称地与二分体位点结合。该方法还能显示二分体中对称地相对突变是否能改变结合。在本章的后面我们还要对这些方法作详尽说明。

2. *二聚体的凝胶迁移变动分析*（EMSA）：这种策略可以用于研究两种蛋白质是否形成异源二聚体或一个蛋白质是否是一个同源二聚体。在同源二聚体的情况下，该方法要求研究者拥有相同蛋白质的一组缺失衍生蛋白，这些衍生蛋白分别与同一位点结合，但在一个 EMSA 中产生具有各自独特迁移率的不同复合物。选择两种这样的衍生蛋白质（通常可通过体外转录和翻译合成这些蛋白质）并在它们处于单独或组合的状态下分析它们对该位点的结合能力。在寡聚体中，单体的结合符合二项式分布 $(a+b)^n$，其中，a 是一种衍生蛋白的浓度，b 是另一种衍生蛋白的浓度，n 是寡聚状态。如果该蛋白质是一种同源二聚体或异源二聚体，则 EMSA 复合物的分布可以通过公式 $a^2+2ab+b^2$ 进行预测。因此，每个衍生蛋白质单独生成一个复合物（如图 11.4B 所示），并且两种衍生蛋白质在一起又形成一个迁移率居中的新复合物。如果 a 和 b 以等摩尔浓度存在，三种复合物则呈现 1∶2∶1 的分布。不论 a 和 b 的浓度是多少，它们都可以代入该公式，预测复合物的分布。如果蛋白质是一种三聚体，EMSA 复合物的分布可根据公式 $(a+b)^3$ 进行预测。这个方法对于同源二聚体和异源二聚体都管用。在两种不同的蛋白质形成稳定异源二聚体的情况下，如 Fos 和 Jun，可能会观察到复杂的 EMSA 条带模式。例如，单独的 Fos 不能在凝胶中迁移，但单独的 Jun 却可以（尽管很弱），而两种蛋白质组合在一起将会产生最有效的复合物。这种分析存在的问题是，如果蛋白质含有一个强有力的二聚化界面，一旦两种单体混合，就会形成异源二聚体，可能不会再发生交换，从而引起该分析的混乱。避免这一问题的一种手段是使用多种盐和缓冲液条件，因为某些条件会促进亚基间的交换。另一个更常见的选择是使用体外转录和翻译共合成两种衍生蛋白。这样，两种合成的 RNA 可以合并在一起，在同一混合物中共翻译，从而使亚基在合成时有机会结合。

图 11.4C 显示的是利用 GAL4 衍生蛋白进行真实实验的例子。单独合成或共翻译两个不同的 GAL4 衍生蛋白，并与一个标记的 17 mer GAL4 位点一起孵育，之后在非变性聚丙烯酰胺凝胶上电泳分离。图 11.4 中也已给出凝胶放射自显影结果。由于它们的大小不同，每个衍生蛋白产生了一个独特迁移的复合物（泳道 1 和泳道 2），而当通过体外翻译一起合成时，这两种衍生蛋白生成一种迁移率居中的新迁移复合物（泳道

3）。这种新的单一复合物的出现是异源二聚体形成的证据，并说明 GAL4 以二聚体的形式与其位点相结合，如放射自显影旁的图解说明。但是，将两种单独合成的衍生蛋白混合在一起却并不能产生迁移率居中的复合物，从而说明在所用条件下 GAL4 蛋白具有强有力的二聚化界面，并且一旦复合物形成，亚基就不能再交换（泳道 4）（见 Carey et al. 1989）。

3. *化学交联及 SDS-凝胶电泳*：戊二醛普遍用于交联实验，以交联伯胺，虽然也有更特异的二价交联剂。不同浓度的重组调控因子与递增量的戊二醛一起进行孵育，然后用伯胺（Tris、乙醇胺等）淬灭，产物与未处理的蛋白质一起在 SDS-聚丙烯酰胺凝胶上电泳分离。蛋白质从其预期分子质量类型到更大分子质量类型的迁移通常意味着更高级寡聚体的形成。然而，戊二醛相对来说是非特异的，因此必须和对照（如 BSA，牛血清白蛋白）联合作用。在用于 EMSA 的蛋白质浓度下，以及 DNA 存在与否的情况下，测量二聚体的形成也是有益的，这样的信息可有助于确定二聚化的机制。

4. *凝胶过滤树脂*：凝胶过滤层析是一种相对过时的方法，该方法可基于分子质量分离单体和二聚体。但由于蛋白质的形状会显著影响凝胶过滤柱中蛋白质的迁移，所以这种方法不太可靠。对分子质量的精确测量依赖于凝胶过滤实验和沉降速度实验的结合。可将复合物在这类实验中的表观分子质量与单体的实际分子质量进行比较，以确定蛋白质的寡聚状态。

5. FRET *分析*：FRET 提供了一种分析异源二聚体形成的手段。一般来说，通过在一个独特位置上放置一个半胱氨酸残基（通常在蛋白质末端）并且之后共价连接一个染料分子，将一个异源二聚体的每个成员进行修饰。例如，可将荧光素连接到一个单体上，而将罗丹明连接到另一个单体上。此外，可将不同的单体分别融合到绿色、黄色或青色荧光蛋白上。当一个染料被特定波长的光所激发时，它会释放出能量，该能量激发另一个染料，使其发射不同波长的荧光，这种荧光可通过荧光分光光度计测量。这样，能量从一个荧光团转移到另外一个荧光团。FRET 的优点是距离依赖性的能量转移，其能量转移的效率与染料间距离的 6 次方相关（Hillisch et al. 2001）。因此，依赖于染料或荧光蛋白间的特定组合，能量传递将在 10～100Å 的范围内发生。在结构已知（如诸如 Fos 和 Jun 的 bZIP 蛋白质）的情况下，将染料放在合适的位置上是相当容易的。在单体状态下不会发生能量转移，而在二聚体状态下能量转移就会发生，这是因为异源二聚体中两个单体相互接近。可用这种方法来确定二聚化的比率及 DNA 对这种比率的影响（Kohier and Schepartz 2001a，b）。

解离常数分析

因为一个识别位点的生物活性（即激活或抑制）有时与该位点对调控蛋白的亲和力成比例，因此亲和力是一个用于比较蛋白质与相关但不同的结合位点结合的有用的标准。蛋白质对其位点的亲和力一般用解离常数或 K_d 表示。K_d是描述反应物与产物比率的平衡常数。一旦将蛋白质和 DNA 混合，只有在两种分子的结合和解离达到恒定速度后，才能达到平衡。因此，在测量 K_d 前，反应必须已经达到稳态。K_d 表示为摩尔浓度，一般来说，K_d 依赖于温度、缓冲液及盐等条件而变化，这种可变性对每个蛋白质都是独一无二的，因此，两种蛋白质 K_d 的绝对测量一般是不可比的，除非它们是在相

同的条件下进行的，最好是在一起测定的（见 Jen-Jacobsen 1997）。DNA 结合蛋白的 K_d 值通过硝酸纤维素滤膜结合、EMSA 或 DNase I 足迹测定，尽管最近的技术进步产生了一种利用表面等离子体共振测量 K_d 的仪器。用于 K_d 测量的表面等离子体共振和其他物理方法，如荧光各向异性，在基因表达领域中越来越普遍，但它们需要特殊设备。

一种常见的误解是蛋白质与 DNA 的比例在 K_d 的确定中是重要参数。除非在非常特殊的情况（下文中评论）下，该比例实际上是没有意义的。相反，蛋白质和 DNA 的绝对浓度，以及研究结合的条件对于该测量才是重要参数。因为 K_d 测量正变得越来越普遍，所以我们讨论一下进行这种分析时需要考虑的一些简单原则。我们以 DNA 位点（S）和激活因子蛋白（P）之间简单平衡的分析开始。该反应如下：激活因子（P）与其 DNA 位点（S）结合，形成激活因子/DNA 复合物（PS）。该反应过程用公式表示为

$$\mathrm{PS} \xleftrightarrow{K_d} \mathrm{P} + \mathrm{S} \tag{1}$$

按照惯例最常用解离反应来说明，该反应的代数式为

$$K_d = \frac{[\mathrm{P}][\mathrm{S}]}{[\mathrm{PS}]} \tag{2}$$

研究人员在研究 DNA-蛋白质相互作用时通常会遇到下面三种情况。

1. 亚饱和：当反应混合物中有数以十亿计的 S 和 P 分子，但二者的浓度都远低于 K_d 时，发生这种情况。在这样的条件下，即使存在个数相等的分子，组分的浓度也不足以形成相当可观的复合物（小于 1%）。在 EMSA 反应中，观察不到迁移率变动的复合物。

2. 单组分饱和：当一个组分的浓度，如［P］，等于或大于 K_d，而［S］是极限浓度时，P 将开始饱和 S。在 EMSA 中，当［P］等于 K_d 时，50%的探针会被饱和，因此 50%的 DNA 变动到结合型。在这种情况下，尽管 50%的 S 被饱和，但仍存在相当一部分的非结合型的 P。

3. 化学计量性相互作用：当两个组分的浓度都超过 K_d 时，它们以化学计量的方式相互作用。因此，如果 P 和 S 以远大于 K_d 的相等浓度出现时，它们将总是以复合物的形式存在，与亚饱和的情况相反。

后两种情况的例子在专题 11.1 中解释。

确定活性蛋白和 K_d 值

几种完善的用于确定 K_d 的方法使用 EMSA 或定量 DNase I 足迹分析。最传统的方法之一依赖于使用方程式（8）中描述的概念：当蛋白质和 DNA 浓度都上升到高于 K_d 时，它们的相互作用就变成化学计量性反应。因此，如果 DNA 浓度已知，就可凭经验确定活性蛋白质浓度。之后，该活性蛋白质浓度可用于滴定，在这些滴定中 DNA 是极限的，以确定 50%的饱和点或 K_d。这种方法依赖于三个条件：①蛋白质必须足够纯，以便通过 Bradford、Lowry 或者缩二脲法等分析方法精确确定其绝对浓度；②获得足量的、带有高亲和力位点的寡核苷酸；③可以精确测量溶液中蛋白质结合的分析方法。

如果采用 EMSA，则应进行预实验，以证明 EMSA 和 DNase I 足迹分析产生相似

的结果。EMSA更容易定量，因为通过放射自显影（即通过磷屏成像仪）的光密度扫描，可以测量结合型和未结合型DNA。有时，EMSA的条件并不总是最佳的，并且在凝胶电泳过程中会有利于蛋白质-DNA复合物的解离。然而，有一种所谓的笼合效应，在这种效应中凝胶孔的有限体积使解离降到最小，或促进迅速的再结合。

第一步是通过以非常少量的^{32}P标记寡核苷酸（约10^{-11}mol/L）滴定反应中的蛋白质，以确定蛋白质的表观K_d。例如，在EMSA中产生50%占据所需的蛋白质浓度被认为是表观K_d。这里使用“表观K_d”这一术语，是因为虽然加入的蛋白质量是已知的，但“活性”蛋白质的量是未知的。由于纯化过程中关键残基（如半胱氨酸残基）部分变性或氧化，许多重组蛋白，特别是那些需要特殊结构（如锌成核的蛋白质折叠）的重组蛋白，都不具有完整的活性。为了保证在最初的滴定中DNA确实是极限的（即低于K_d），应将DNA的量增加或降低5倍，并且占据分数（fractional occupancy）（50%）应该保持相同。方程式（4）和方程式（6）说明了这一概念。

专题11.1 K_d情况的例子

对于只有一种组分是饱和的情况，让我们假设P对其位点的K_d是10^{-9} mol/L。当我们将总蛋白浓度[P]设置为等于K_d（10^{-9} mol/L)，并将总位点浓度[S]设置为以下两个小于K_d的浓度：(A) 2×10^{-11} mol/L和(B) 2×10^{-15} mol/L时，现在让我们测量S的占据率。因为K_d是常数，所以必须使用近似法运算DNA和蛋白质的浓度，以达成与K_d相等的最终计数。为了尽量减少运算，已经设定好蛋白质和DNA位点的用量。在情况A中，将会得到如下结果。

$$K_d = 10^{-9} = \frac{(10^{-9})(10^{-11})}{(10^{-11})} = \frac{[P][S]}{[PS]} \tag{3}$$

（请注意：方程式（3）是方程式（2）中代入数值的结果。

请注意：与S（10^{-11}）结合的P的量是可以忽略不计的，该量不能有效地改变游离型蛋白质的浓度（10^{-9}mol/L)，因此可以使用游离型蛋白质的浓度近似值（即10^{-9}）代替代数。游离型蛋白质的确切浓度可以利用代数进行计算，但是对总体结果几乎没有影响。还要注意的是，我们以2×10^{-11}mol/L的位点浓度开始，该浓度分割在S和PS之间，即

$$\frac{[S]}{[PS]} = \frac{(10^{-11})}{(10^{-11})} = 1 = 50\%\ \text{DNA 占据率。} \tag{4}$$

因为结合型DNA（PS）和游离型DNA（S）的量相等[方程式（4)]，所以我们说此时有50%的DNA被占据。因此，当DNA浓度是极限的（即远低于K_d），50% DNA被结合的游离型蛋白的浓度就是K_d的操作性定义。结合型或复合型DNA的百分比通常被称为饱和分数（fractional saturation）或占据率。

现在，我们检查情况B。很明显，如果我们将DNA浓度调整为2×10^{-15} mol/L，则

$$10^{-9} = \frac{(10^{-9})(10^{-15})}{(10^{-15})} = \frac{[P][S]}{[PS]} \tag{5}$$

我们可以明显看出 DNA 的占据分数与当 DNA 是 2×10^{-11} mol/L 时相同。因为与 DNA 相结合的蛋白质的浓度可以忽略不计，结合型和未结合型 DNA 的比例仍是 1，相当于 50%的占据率，即

$$\frac{[S]}{[PS]}=\frac{(10^{-15})}{(10^{-15})}=1=50\%\ \text{DNA 占据率} \tag{6}$$

有关要点如下：①当蛋白质浓度等于 K_d 时，其浓度可能大于 DNA 位点浓度数千倍，然而仅有部分 DNA 被占据；② 当 DNA 浓度远低于 K_d 时，在任何给定的蛋白质浓度下，都将观察到该位点的相同占据分数。

使用这样一组条件，其中 DNA 浓度低于 K_d，我们可以将 [S] / [PS]（右轴）或 DNA 探针的饱和分数（左轴）绘制为表达在对数刻度上的蛋白质浓度的函数。图 11.6 显示饱和度随蛋白质浓度的增加呈抛物线增加。注意，在前面所述的简单双分子相互作用中，为了将探针饱和度从 10%提高到 90%，或将 [S] / [PS] 从 9 减小到 0.11，蛋白质浓度需增加为原来的 81 倍。

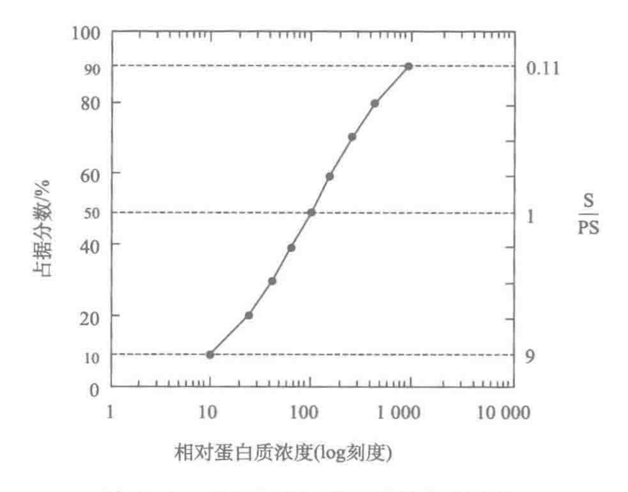


图 11.6　简单双分子反应的结合等温线。

当两者的浓度都等于或高于 K_d 时，就会观察到不同的结果。如果加在反应体系中的总蛋白质浓度仍为 10^{-9}，而将 DNA 浓度调整至 10^{-7}，用近似法可以得到下列结果，即

$$K_d=10^{-9}=\frac{(10^{-11})(10^{-7})}{(10^{-9})}=\frac{[P][S]}{[PS]} \tag{7}$$

（相反，也可以提高蛋白质浓度，K_d 并非指反应中的任一组分）。

其结果是，几乎所有的起始蛋白质进入与 DNA 形成的复合物，但是总的游离 DNA 浓度（10^{-7}）并没有改变，因为它远远高于蛋白质的浓度。考虑另一个例子，如果将加入到反应的总蛋白和 DNA 的浓度都提高到远高于 K_d，分别达到 10^{-5}，则

$$K_d=10^{-9}=\frac{(10^{-7})(10^{-7})}{(10^{-5})}=\frac{[P][S]}{[PS]} \tag{8}$$

为平衡P、S和PS的量，几乎所有蛋白质与DNA都必须以化学计量的方式相互作用。因此，几乎所有加入的10^{-5} mol/L DNA和蛋白质都形成PS复合物，较少量（使用近似法为10^{-7} mol/L）以游离型P和游离型S存在。虽然通常认为这后一种情况在标准的DNA-蛋白质相互作用研究中会发生，但通常一种组分是极限的，并且适用方程式（3）和方程式（5）中描述的情况。然而，正如文中所述，这种情况是确定DNA结合蛋白活性的常用方式。

一旦表观K_d已知，就需要将蛋白质浓度提高至表观K_d的50～100倍，甚至更高，这将导致几乎所有的^{32}P标记的寡核苷酸被占据。虽然活性蛋白的精确数量是未知的，但活性蛋白的水平已提高到远高于K_d并进入与DNA以化学计量方式结合的水平范围。下一步就是要逐步增加一个已知浓度的未标记的竞争性DNA位点（以寡核苷酸的形式），以确定反应中需要多少量的DNA才能结合所有的蛋白质，因为这使蛋白质活性分数（fractional activity）的估计成为可能。随着竞争性寡核苷酸用量的提高，它开始与^{32}P标记寡核苷酸竞争蛋白质。因为当未标记的寡核苷酸开始竞争时（即减少EMSA中的占据），DNA浓度超过K_d，推测它必定与蛋白质以化学计量方式相互作用。最终，溶液中所有的活性蛋白质定量地被该寡核苷酸扣押，并且结合的寡核苷酸的量与活性蛋白质的量相同。

当竞争DNA的量开始超过活性蛋白的量时，通过EMSA或足迹法中可观测到未结合的放射性标记DNA。当未结合寡核苷酸竞争了50%的放射性标记寡核苷酸时，竞争寡核苷酸超过活性蛋白的量两倍（即活性蛋白的浓度相当于寡核苷酸浓度的一半）。由于寡核苷酸的摩尔浓度是已知的，便可以容易地计算出活性蛋白的浓度。一旦活性蛋白的浓度已知，就可回到原来的实验并将表观K_d替换成真正的K_d。这种方法非常简便，并且可不做任何特殊的代数运算。这种方法的优点是，如果目的蛋白是一个寡聚体（如二聚体），在化学计量学方法中使用高浓度的蛋白质将有利于形成二聚体类型的分子。这样，可确定反应混合物中有活性的二聚体的量。注意，因为根据K_d方程，游离DNA总是与结合的DNA相关，可用寡核苷酸竞争来计算在任何蛋白质浓度下的K_d。然而，代数运算更加复杂，并依赖于该方法作出几种关于蛋白质活性或寡聚化状态的假设。

DNA-蛋白质相互作用的分析及建模

幸运的是，由于在DNA结合蛋白及其与DNA形成的复合物晶体结构或核磁共振（NMR）结构的解析中取得的迅速进展，使我们对DNA识别的了解已相当深入。一般来说，如果知道一种蛋白质属于一种可识别蛋白家族，而该蛋白家族的晶体结构已获得，则可以轻易地对调控因子的序列偏好性提供一个总体的看法，也可以对理解该蛋白质如何识别DNA提供一个概要。对于那些属于bZIP、bHLH、同源结构域及锌指家族的蛋白质，这一点尤其真实。目前已有从蛋白质序列中搜索HTH模体的网站，如http://predictdnahth.rutgers.edu/和http://www.ebi.ac.uk/thornton-srv/databases/

DNA-motifs/。

核酸酶和化学探测可为蛋白质如何识别其位点提供宝贵信息，这种探测可以用来完善模型，测试晶体结构或 NMR 结构的有效性，以及探查以前尚未表征的调控蛋白家族的结合。在下面的几节中，我们将讨论通过这样的一些方法可获得的信息。我们首先讨论如果研究人员有一种 DNA 结合蛋白，但高亲和力识别位点却未知时，鉴定位点所必需的方法。

高亲和力 DNA 识别位点的鉴定

很多情况下，研究者手头都有一种蛋白质，因为该蛋白质分离自鉴定重要调控蛋白的遗传筛选或可能被评为原癌基因。可能有一些该蛋白质与 DNA 结合的证据，但其作用位点不清，诸如 Myb、Myc、MyoD 和孤儿核受体的蛋白质都是这种情况的典型例子，在这种情况中，为了鉴定实际的作用位点都需花费一段时间。在这样的一些情况下，怎样才能鉴定位点从而启动其 DNA 结合机制的研究呢?

在酿酒酵母（*Saccharomyces cerevisiae*）和果蝇（*Drosophila melanogaster*）这些遗传学研究容易进行的物种中，通过遗传学筛选已经克隆了许多调控蛋白，并且通过受影响基因的启动子突变已经鉴定并确证了它们的启动子位点。通过遗传学方法使位点或因子失活，可以判断该因子在体内是否通过该位点发挥作用。此外，可用染色质免疫沉淀（ChIP）分析确认该蛋白质在体内与预测位点结合。在哺乳动物系统中的研究，一般依赖于大量的方法，这些方法将突变与转染、ChIP 或生化分析相结合。下面所列的策略描述了用于高亲和力位点鉴定的方法，该位点是启动结合机制研究所必需的，但这些方法并不描述如何确定该位点是否是一个与天然应答启动子相关的生理性识别序列（这个问题在第 9 章中已详细讨论）。

鉴定 DNA 位点的方法

有很多用于鉴定目的因子结合位点的方法。有些方法旨在产生一个用于基本生化分析的位点，而其他方法则用于一些全基因组方式，以鉴定启动子和增强子中的天然位点。全基因组方法的数据集可以被注释，并与基因表达模式进行比较，以破译调控网络。在可预见的未来，大部分转录因子中几乎所有的结合位点都将被鉴定。在这个时候，重要的将是确定增强体的概念是否可以被推广普及和（或）位点的特定组合是否界定共激活因子和共抑制因子募集有关的独特功能。

启动子扫描：该方法涉及利用重叠的限制性片段或人工合成的寡核苷酸为探针，以及纯化的重组蛋白或细胞提取物为结合蛋白来源，通过 DNase I 足迹和（或）EMSA 分析对应答基因的启动子进行扫描。通常通过使用启动子或增强子的系统发生分析可极大地促进该方法的应用。

选择技术：这些方法可以分为几类，是在结合位点未知的情况下，为简单生化研究鉴定高亲和力位点的首选方法。在一种形式的选择技术中，将 DNA 结合蛋白连接到固相亲和基质上，使随机 DNA 的混合物（寡核苷酸、随机限制性片段或含有应答性启动子的片段）通过这种亲和基质。在基质中保留下来的片段可以再一次选择，然后亚克隆用于生化分析（Kalionis and O'Farrell 1993）。该项技术的一种形式是“选择并扩增结合位点（SAAB)”分析（专题 11.2）（Blackwell 1995）。SAAB 方法采用

一种如固相亲和基质或 EMSA 的结合技术，鉴定与蛋白质相结合的 DNA 寡核苷酸。然后对 DNA 片段进行纯化和聚合酶链反应（PCR）扩增，并再次选择。通过改变结合反应的严紧性，就可以得到用于结合分析的高亲和力位点。该方法的另一种形式是“指数富集配体系统进化（SELEX）”，该技术一般使用寡核苷酸。然而，通过向亲和介质中加入片段化的基因组 DNA 并用盐洗脱，也可以使基因组 DNA 片段适用于 SELEX。因为少量的 DNA 难以观察，故可以将接头连接到筛选 DNA 的两端并进行 PCR 扩增。经过几轮筛选和 PCR 扩增后，将基因组位点克隆，然后通过测序或传统方法（如 EMSA）进行分析，将 DNA 与微阵列杂交来确定基因组中的精确定位也是可行的（Shostak et al. 2004）。

染色质免疫沉淀：这种方法依赖于免疫沉淀体内形成的 DNA-蛋白质复合物的能力。在最初的方法中，标准的 ChIP 实验在戊二醛处理的细胞上进行，但在免疫沉淀后，将基因组 DNA 从蛋白质上解离并克隆（Weinmann and Farnham 2002）。最近的也是流行的方法是使用微阵列（ChIP-Chip）或高通量 DNA 测序技术（ChIP-Seq）。在 ChIP-Chip 中，免疫沉淀后，将 DNA 连接接头，并用引物和荧光标记 dNTP 扩增。之后，扩增的 DNA 与涵盖启动子和增强子的叠连微阵列杂交。最新的阵列涵盖大多数已知的启动子，有的也包含增强子（Kim and Ren 2006）。在 ChIP-Seq 中，免疫沉淀的 DNA 连接接头，变性，然后加入到一个流动室中，该流动室由一个光学透明的面构成，在该光学透明面上固定了密集的与接头互补的引物。之后，DNA 在流动室中扩增，形成高密度的扩增 DNA 簇（大于 5×10^7），每一簇都代表一个来自 ChIP 反应的不同 DNA。这些 DNA 随后在流动室中使用荧光标记的 dNTP 进行测序。ChIP-Chip 的检测偏向于已知启动子或增强子，与 ChIP-Chip 不同，ChIP-Seq 是一种无偏的全基因组技术（Robertson et al. 2007）。理想的情况下，通过比较 ChIP-Chip 和 ChIP-Seq 两种方法得到的序列，原理上便可以得到目的位点的一致序列。然而，能否容易地获得位点将取决于叠连阵列上探针的大小和密度（ChIP-Chip）或者测序片段的大小（ChIP-Seq）。值得注意的是，这些技术一般用于聚焦于特定转录因子或染色质修饰的全基因组分析（见第 8 章和第 9 章）。如果一个因子已经进行了这类研究，建议读者参考已发表论文中的数据，以获得合适的位点。

蛋白质结合微阵列：这种方法涉及制备一个荧光标记的转录因子或有荧光标记抗体的表位标记的因子。在该方法的一种变种中，将蛋白质结合到含有所有可能的 10 mer 序列的双链 DNA 微阵列上（Berger et al. 2006）。结合后阵列上出现各种各样的荧光点；通过将荧光强度赋予该位置代表的 DNA 序列，便可估计出蛋白质与不同位点的相对亲和力（Bulyk 2006）。从原理上来说，任何一套 DNA 片段都可以用于制备微阵列。例如，代表酵母基因间区域的 PCR 片段可用于制备阵列，以研究一组酵母转录因子的结合位点（Mukherjee et al. 2004）。尽管有可能鉴定任何转录因子的结合位点谱，但该技术需要高度专业的知识和设备。之后，结合位点的数据可用于生化研究，也可以用于计算分析，以鉴定基因上游区域，并将其与基因表达及调控信息联系起来（Warner et al. 2008）。关于现行技术的更新及细节，请参考 Martha Bulyk 的网站 http://the_brain.bwh.harvard.edu/。

专题 11.2　SAAB 分析

SAAB 方法（图 11.7）由 Blackwell 和 Weintraub（Blackwell 1995）创立，以用于比较 bHLH 蛋白 MyoD 和 E2A 的同源及异源二聚体的结合位点。Elhngton 和 Szostak（1990）也创立了独立的但概念上相似的方法，以筛选具有独特结合性质的 RNA 分子。尤其是应用于 RNA 时，该方法在随后的出版物中被称为 SELEX（Tuerk et al. 1992）。SAAB 方法至今已用于鉴定数十种调控蛋白的结合位点清单。

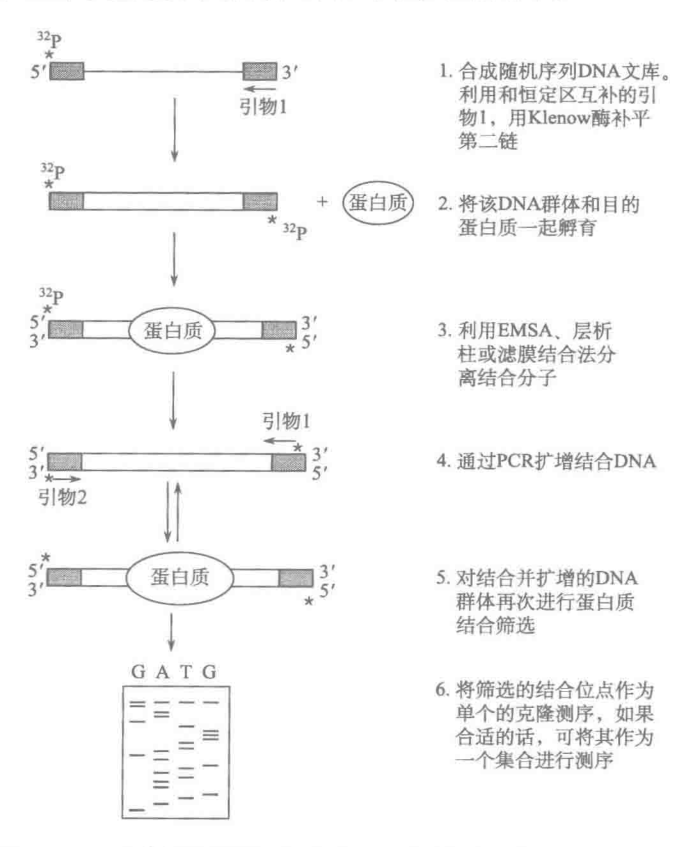


图 11.7　SAAB 分析流程图（经允许改编自 Blackwell 1995，© Elsevier）。

该方法采用与 EMSA 相偶联的 PCR 技术，富有创意，且简单易行。首先构建含有随机序列及其两侧具有恒定 PCR 位点的寡核苷酸，并进行^{32}P 末端标记。合成与恒定区互补的第二个寡核苷酸，并将其与第一个寡核苷酸退火。之后，将退火后的寡核苷酸用作 Klenow DNA 聚合酶的引物，用 Klenow DNA 聚合酶填平缺口，以产生双链分子。然后将^{32}P 标记的分子与饱和或亚饱和浓度的蛋白质一起孵育，通过凝胶迁移实验从未结合的 DNA 中分离出复合物。通过 PCR 再次扩增含有被选择序列的结合型 DNA，并进行一次或多次重选步骤。该方法的一个优点是能够控制结合条件和选择循环数，以区分高亲和力位点和低亲和力位点。

需要注意的是选择技术不一定非要用 EMSA。DNA 在诸如连接谷胱甘肽-S-转移酶(GST) -转录因子融合蛋白的谷胱甘肽-琼脂糖固相亲和介质上的层析，或以转录因子抗体进行的 DNA-蛋白质复合物的免疫沉淀，都是鉴定结合位点的替代方法。有些研究将 SAAB/SELEX 应用于 GST 标记的转录因子，以鉴定含有调节因子结合位点的基因组 DNA 片段 (Shostak et al. 2004)。

DNA 片段的准备及基本理论

所有核酸水解的和化学的探测方法都基于用于描述 DNA 测序的套嵌组理论（见 Maxam and Gilbert 1977)。一个套嵌组含有一个独特末端和一个可变末端（图 11.8)。当含有一个独特末端（被一端标记，通常是^{32}P）和一个可变末端（代表修饰或切割位点）的 DNA 分子通过变性凝胶电泳分级分离时，在凝胶上测得的末端标记的单链 DNA 片段的大小，就是到可变端的距离的精确测量（即可变末端相对于独特末端的位置)。测序及酶和化学探测的原理是产生含有独特末端和定位于特定碱基的可变末端的 DNA 片段。凝胶上放射性单链 DNA 分子的长度指定了该特定碱基的确切位置。在大多数足迹方法中，通过对 DNA 片段一个末端的直接^{32}P 末端标记产生独特末端，其中 DNA 片段既可以使用限制性内切核酸酶酶切的 DNA 片段，也可以使用 PCR 扩增的 DNA 片段。

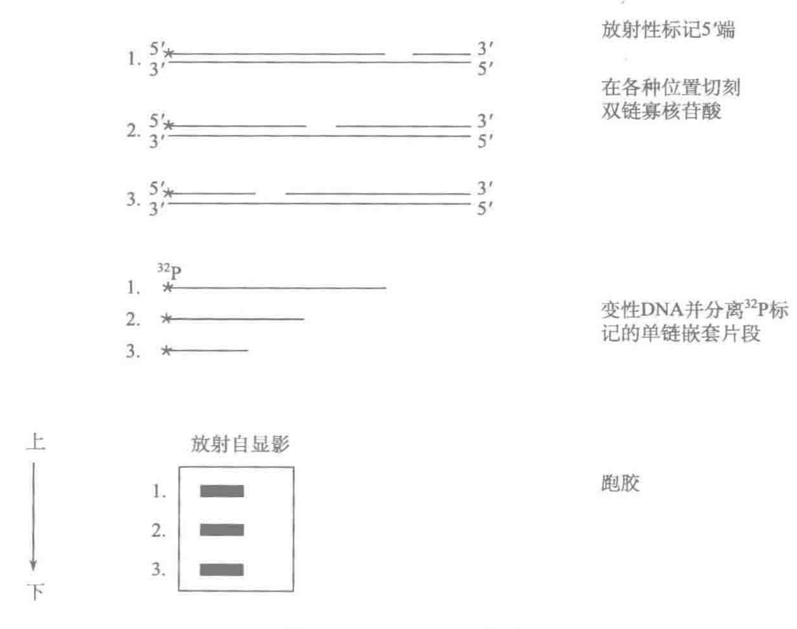


图 11.8 嵌套寡核苷酸。

在启动任何一个DNA-蛋白质相互作用的分析中，应从简单的方法开始。将重组蛋白纯化后，产生带有目的位点的DNA片段。最好使用DNA小片段，这些片段的带型在测序胶上能分辨到碱基对。虽然合成的寡核苷酸通常足以在EMSA中测量结合，但由于下述原因，它们只适合研究小蛋白质对短位点的结合。因此，通常先将结合位点亚克隆到操作简单的载体上。对于7～17 bp范围内的位点，该位点最好位于一个50 bp区域的中间，这样的50 bp片段既可以从凝胶上切出，作为独特的^{32}P末端标记限制性片段，也可以用其中一个引物被^{32}P标记的引物对通过PCR扩增制备。经过片段纯化和随后的蛋白质-DNA复合物的酶或化学分析，将产物在7%～10%的聚丙烯酰胺凝胶上分级分离。该操作会产生详细的切割梯带，足以进行仔细而定量的分析。

用于评价结合的酶探针

对蛋白质与其位点结合提供广泛指示的两个主要方法是DNase I足迹（Galas and Schmitz 1978）和外切核酸酶III足迹（专题11.3）。由于核酸酶的分子大，以及DNA结合蛋白的空间位阻，产生的足迹通常比识别位点本身要大得多（一般要大8 bp或更多），难以定位结合位点的精确边缘。但是这些方法为任何DNA-蛋白质相互作用的研究和建模提供了一个必要的起点。还有几个利用小分子化合物的方法，这些小分子化合物，如菲咯啉铜（即*O*-菲咯啉-CuSO4；OP-Cu）或甲啶丙基EDTA（MPE），可结合在DNA小沟中并切割DNA（专题11.3）。这些试剂一般能更精确地确定结合位点的边界（Sigman et al. 1991）。

小沟及DNA骨架化学探针

蛋白质沿DNA骨架的定位可以提供与其识别机制相关的重要信息。如前所述，蛋白质与磷酸骨架的相互作用在两个方面发挥着重要作用，一是提供结合能量，二是使这些相互作用定位到大沟中的蛋白质识别元件上。用于探测小沟和骨架相互作用的技术主要有两种（专题11.4）：羟自由基保护法（方案11.2）和乙基化干扰法（方案11.3）。DNase I和OP-Cu足迹也能结合于小沟，但只提供大沟和小沟相互作用较为宏观的信息。与这两种方法不同的是，羟自由基保护法（因为自由基分子小，易扩散；Balasubramanian et al. 1998）和乙基化干扰法可专门提供更详细的糖/磷酸骨架相互作用的信息。由于这种原因，这些技术在足迹分析中已变得非常流行。羟自由基足迹法鉴定靠近糖残基的相互作用，而乙基化干扰法鉴定与磷酸发生的相互作用。

专题11.3 DNase I足迹和外切核酸酶足迹

DNase I足迹

DNase I足迹（图11.9）是研究调控蛋白与其识别序列结合的最广泛使用的方法之一（见方案11.1）（Galas and Schmitz 1978）。DNase I作用于DNA的小沟，虽然它以一种相对非特异性的方式结合并切割小沟，但其活性依赖于其结合位点采取的某种基于序列的结构特征（Suck 1994）。位点中局部结构的变形，如螺旋桨式扭转或局部小沟宽度的差异，都会显著影响DNase I的结合和（或）切割特性。

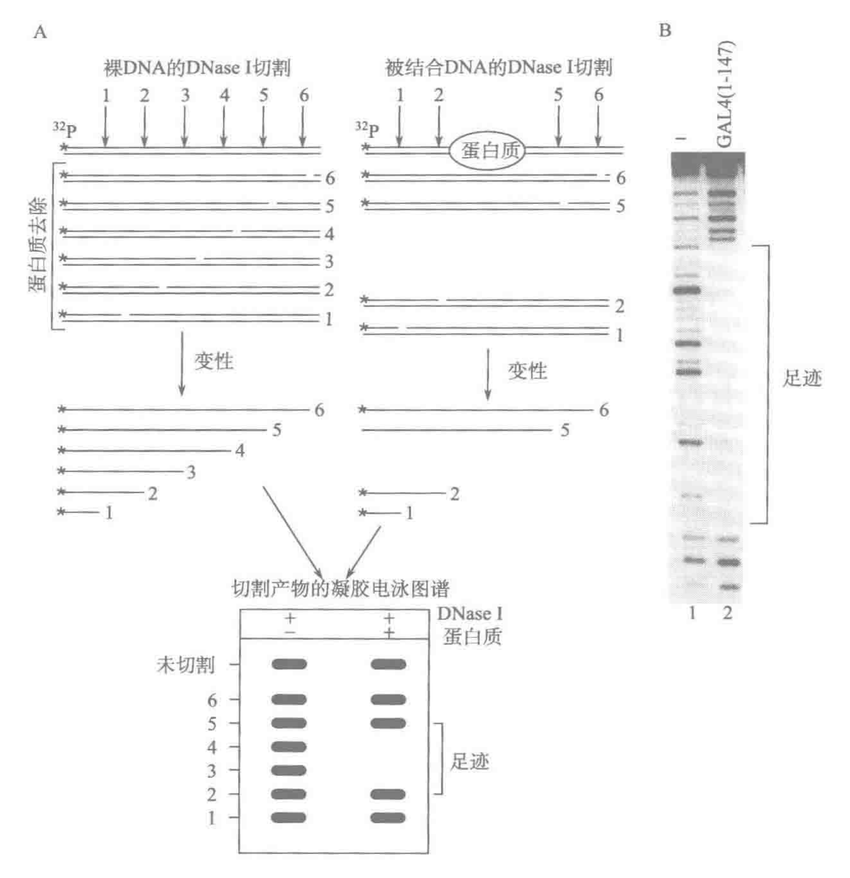


图 11.9 (A) DNase I足迹原理。(B) GAL4 DNA结合结构域分析示例 (B经允许改编自 Carey et al. 1989, © Academic Press Ltd.)

在 Ca^{2+} 和 Mg^{2+} 存在的情况下，DNase I 切割双链 DNA 中的一条链（镁离子能引起双链切割)，并产生 5′-PO_4 和 3′-OH 产物。结合在 DNA 上的蛋白质会保护 DNA 位点免受切割。对于许多蛋白质来说，DNase I 足迹明显大于识别序列（可能是因为 DNase I 蛋白本身很大)，因此 DNase I 被视为低分辨率探针。因为 DNase I 的形状不对称，且在特定的 DNA 链上会更靠近位点边缘的其中一端，所以产生的足迹经常是错开的。图 11.9A 为足迹法的原理图，图 11.9B 显示了使用 GAL4 DNA 结合域的研究实例。在足迹法中，相对重要的是平均每个 DNA 分子只切割一次，以便观察到消化梯带，这样就只需要探针的部分切割。在蛋白质结合的地方，DNase I 不会切割，从而产生丢失的切割产物，这些丢失的切割产物在凝胶便形成“足迹”。

外切核酸酶 III

外切核酸酶 III (ExoIII) 是一种从大肠杆菌中分离出来的分子质量为 28 kDa 的核酸酶。它优先结合双链 DNA 上的游离 3′ 羟基，然后以半进行性的方式向内切割。它更喜欢与凹进的或平末端的 3′-羟基结合，切割 DNA 释放核苷单磷酸，且具有 3′-磷酸酶活性。ExoIII 和单链内切核酸酶联合使用，产生基因或其启动子区的缺失体。然而，

它也可以用于定位蛋白质上的结合位点（Siebenlist et al. 1980）。当遇到 DNA 位点且该位点被蛋白质结合时，ExoIII 便停止切割（图 11.10）。一般情况下，如果给予足够的时间，核酸酶会等到蛋白质从其位点上脱落，可通过该位点造成放射自显影上的“通读”。ExoIII 足迹比 DNase I 足迹稍小一些。对于 GAL4 蛋白，ExoIII 生成一个 21 bp 的足迹，而用 DNase 时，足迹是 27 bp（Carey et al. 1989）。

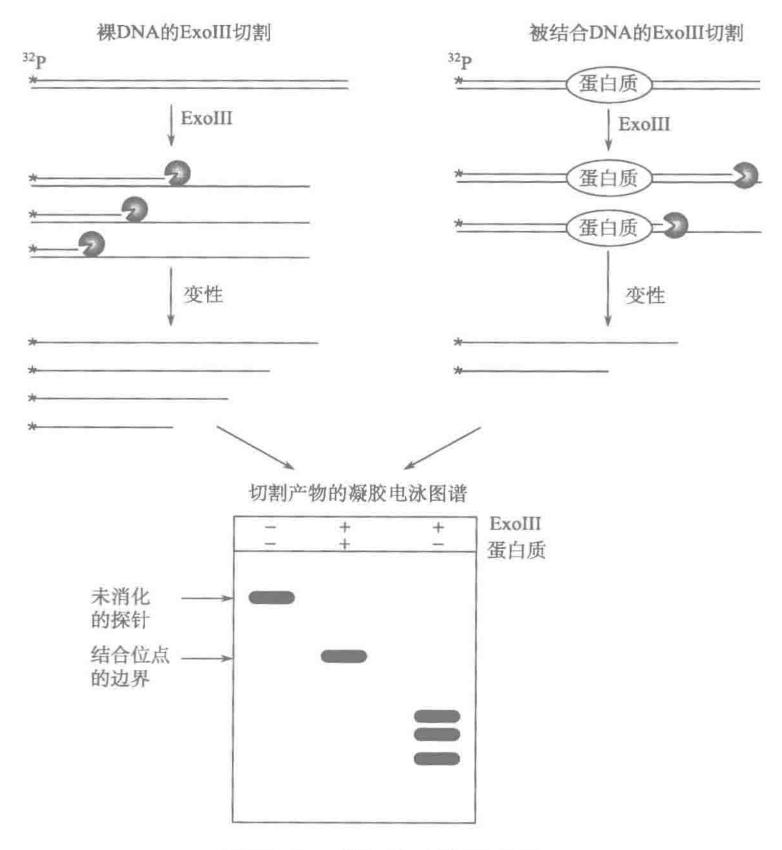


图 11.10　外切核酸酶 III 足迹。

其他通用足迹试剂

虽然 DNase I 和 ExoIII 足迹为位点结合提供了一个宽泛的指示，但更小的化学切割试剂，如甲啶丙基 EDTA（MPE）和菲咯啉铜（OP-Cu）能更精确地确定位点边缘，还能提供关于相互作用的重要结构信息。这两种试剂都结合在小沟中，MPE 插入碱基对之间，它可产生可扩散的羟自由基，这些羟自由基在糖/磷酸骨架上切割 DNA。OP-Cu 被认为可以通过攻击糖上的 C-1 氢原子导致更加定向的切割。OP-Cu 可能是一种特别有用的试剂，因为当用其他方法造成蛋白质-核酸复合物不稳定时，可以在丙烯酰胺凝胶中进行切刻分析（nicking assay）（Sigman et al. 1991）。

专题 11.4 用于小沟相互作用的化学探针

羟自由基足迹

羟自由基可以攻击许多脱氧核糖上的残基，其中对脱氧核糖上的氢原子有偏好（C5′ > C4′ > C3′ = C2′ = C1′），并抽取连接在这些残基上的氢原子（见 Dixon et al. 1991；Balasubramanian et al. 1998）。以前认为它是一种小沟试剂，但更严格地讲，它是探测在糖/磷酸骨架附近的接触的探针（Balasubramanian et al. 1998）。氢原子的抽取引发一系列电子转移反应，最终引起 DNA 链的断裂。羟自由基分子与水分子一样小，不会受到与 DNase I 类似的空间限制。因此，羟自由基足迹法能揭示有关沿骨架初级接触的详细信息。羟自由基对骨架结构中的微小扰动有点不敏感，除非在扭结和弯曲的情况下，切割效率才会改变。不幸的是，羟自由基也攻击蛋白质。有些蛋白质比其他蛋白质对自由基或产生自由基的化合物更敏感。羟自由基通常使用芬顿反应产生（图 11.11A）。图 11.11B 显示典型的羟自由基足迹反应，其中浓度递增的 GAL4 DNA 结合结构域（泳道 3～5）和 DNA 一起孵育，并用自由基切割。当和单独的 DNA（泳

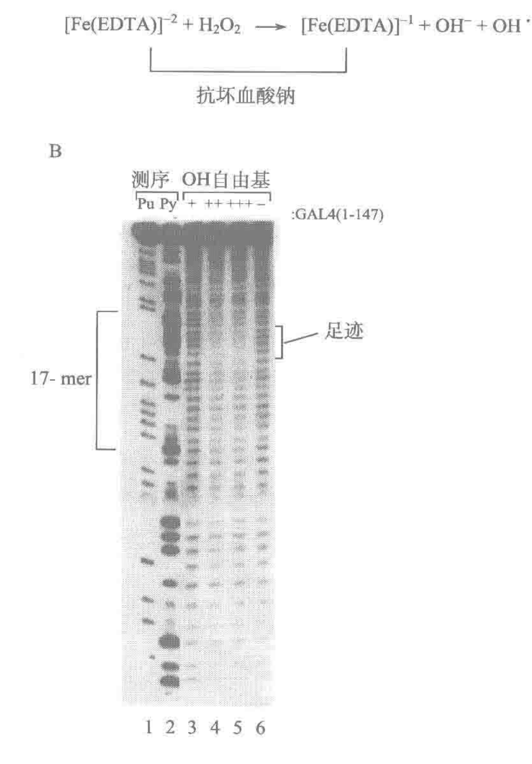


图 11.11 （A）芬顿反应。（B）GAL4 的羟自由基足迹。（B 经允许改编自 Carey et al. 1989，© Academic Press Ltd.）

道 6）相比时，可观察到位于 17-mer 中央附近的一个小足迹。通过将该位置与 Maxam-Gilbert 嘌呤和嘧啶切割测序梯带（泳道 1 和 2）相比，可确定该足迹的位置。这一信息可用于 GAL4 与其位点相互作用的建模（见下文）。

乙基化干扰

乙基化干扰依赖于乙基化磷酸干扰蛋白质结合的能力（Siebenlist et al. 1980；Manfield and Stockley 1994），信息内容与羟自由基足迹相似但不相同。首先，以乙基亚硝基脲乙基化 DNA，乙基亚硝基脲主要修饰沿着骨架的磷酸（图 11.12）。拿捏好修饰进行的程度，要确保每个 DNA 分子一个乙基（图 11.13）。考虑到修饰过程的泊松分布，在相当数量的分子被两次或多次乙基化修饰之前，应该允许只有大约 10%的分子被乙基化（图 11.13）。通过修饰 DNA 直到约 10%的起始探针被转变成梯状条带，可以确定这种火候。关键的是只修饰一次，以致在最后的分析中，很显然被检测的修饰（最靠近 ^{32}P 标记末端的修饰）就是造成干扰的修饰。同一分子上存在过多的修饰会搞乱该分析，这一原理对于其他的干扰和保护分析同样适用。

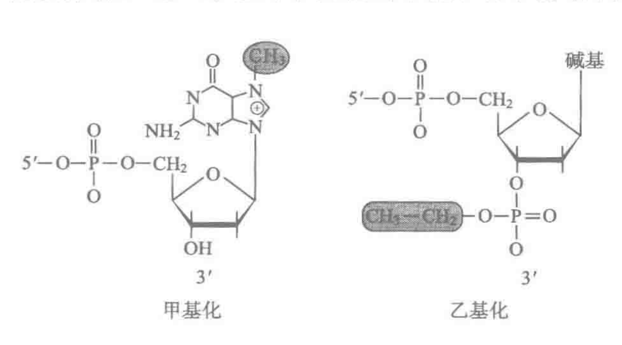


图 11.12 甲基化的 G 和乙基化的磷酸。

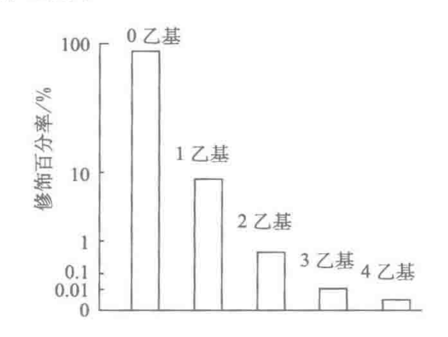


图 11.13 不同乙基状态中 DNA 分子的最佳泊松分布。

修饰后，接近饱和量的蛋白质与 DNA 相结合（图 11.14A）。这里使用接近饱和的蛋白质量是因为虽然修饰减少结合，但并不总是废除结合；高浓度的蛋白质可以克服修饰的有害影响，然后通过 EMSA 分离未结合级分和结合级分的 DNA。结合级分的 DNA 富含在不干扰蛋白质结合的位置上被修饰、并且因此修饰位置不靠近蛋白质的 DNA 分子（图 11.14，步骤 2 和步骤 3）。而未结合级分的 DNA 被修饰在干扰蛋白质结合的位置上。通过首先分离结合级分和未结合级分的 DNA（如来自 EMSA 的条带，步骤 3），再用哌啶处理 DNA（在受影响残基处切割），最后以变性凝胶电泳同步分级分离切割片段与同一片段的测序梯带（步骤 4 和步骤 5），来鉴定修饰点（图 11.14A）。可用 Maxam-Gilbert 化学法产生测序梯带，或者通过直接制备与 ^{32}P 标记分子具有相同 5′ 端的测序引物，再利用双脱氧测序法产生测序梯带。

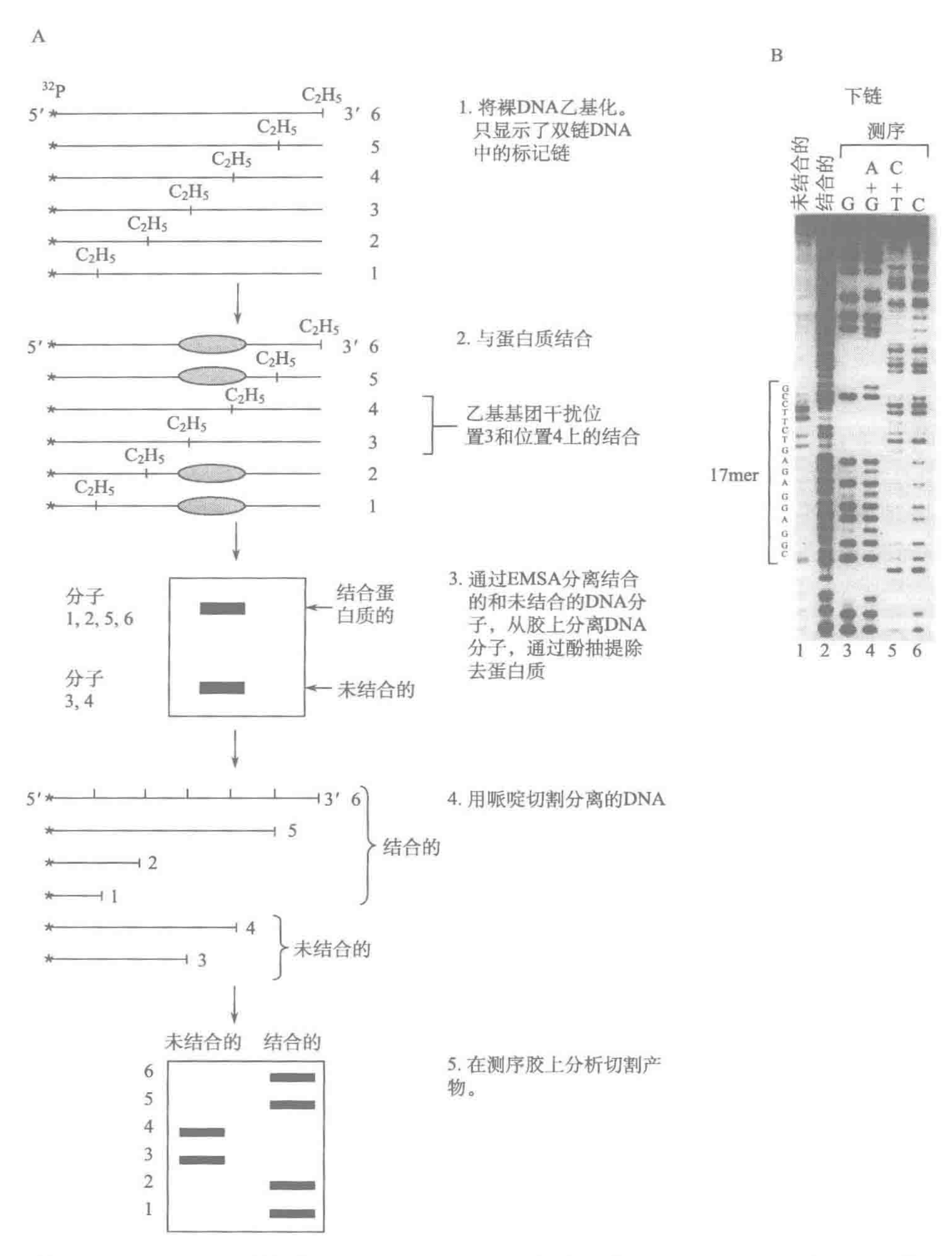


图 11.14 (A) 乙基化干扰。(B) GAL4 实验（B经允许改编自 Carey et al. 1989，© Academic Press）。

该数据通常很清楚，因为在一段给定的DNA上大多数位置都不会产生干扰，因此结合级分含有跨越凝胶长度的梯带。然而少数条带会在结合级分中耗尽，但在未结合级分中富集，这些条带代表干扰结合的修饰及靠近结合蛋白质的位置。一般通过将接触叠加到图示的B-DNA螺旋上，并对这些数据进行概括。如果蛋白质属于结构已知的蛋白

质家族，就可以利用该结构作为出发点对该数据进行建模。

图 11.14B 举例说明了一个使用 GAL4 的典型实验。未结合泳道中的条带表示干扰 GAL4 结合的磷酸，而结合泳道中的条带表示不干扰 GAL4 结合的磷酸。要注意结合泳道中磷酸的缺失和未结合泳道中磷酸的富集。最后通过 Maxam-Gilbert 测序梯带确定位置。这一信息可以用于 GAL4 与其位点相互作用的建模（见下文）。

大沟化学探针

有几种了解蛋白质在 DNA 大沟中如何相互作用的方法，这些方法包括硫酸二甲酯（DMS）保护和干扰法（Siebenlist et al. 1980；Wissmann and Hillen 1991）、核苷缺失法（Hayes and Tullius 1989；Dixon et al. 1991），以及标准的定点突变法。

DMS 方法：DMS 保护的原理与 DNase I 足迹相似，但 DMS 是很小的试剂，因此可以提供有关相互作用的详细信息。DMS 甲基化 G（大沟）环氮中的 N7（见图 11.12）和 A（小沟）的 N3。与 DNase I 产生覆盖整个结合位点的保护不同，只有非常靠近蛋白质的碱基才能被保护，从而免受 DMS 修饰。而且，由晶体或 NMR 结构推断，免受甲基化修饰的碱基经常与直接参与同蛋白质相互作用的碱基相关。但是，由于蛋白质结合可以抑制位点上不接触碱基的甲基化，所以并不总是能表现出这种相关。因此，化学探测分析经常与蛋白质及位点的突变分析相关。下面的例子显示出这些相关性的限度。有些蛋白质对 DMS 保护所需的条件或 DMS 本身敏感，DMS 经常修饰蛋白质，使它们的结合能力下降（如 TFIIIA；Fairall et al. 1986）。这种情况下，可以选择 DMS 干扰分析。

在 DMS 干扰分析中，DNA 在蛋白质不存在时被甲基化，平均每个分子上发生一次甲基化，然后分离被修饰的 DNA。之后，修饰的 DNA 和蛋白质相互作用，分离并分析结合样品和未结合样品。这种分析的方法学和 DNA 处理化学与乙基化干扰法非常相似。为了避免 DMS 保护法可能的复杂化，许多研究者都直接进行 DMS 干扰分析（方案 11.4）。

核苷缺失和诱变：在可用于研究大沟相互作用的其他方法中，核苷缺失和诱变最为简单。在诱变分析中，位点中一个或多个碱基为其他碱基所置换时可以检测到亲和力的改变（通常是下降）。某些碱基对亲和力起主要作用。改变这些碱基将去除对相互作用关键的化学基团，从而导致亲和力下降。

核苷缺失法有些类似于干扰实验（Dixon et al. 1991）。在与蛋白质相互作用之前用羟自由基去除一个核苷（碱基-糖）之后，将有缺口的 DNA 与饱和量的蛋白质一起孵育，通过 EMSA 分离结合级分和未结合级分的 DNA。将条带切下来，在凝胶上与测序梯带一起电泳。结合级分富含缺失不重要碱基的分子，而未结合级分富含缺失对相互作用重要的碱基的分子。该方法提供的信息量可能比 DMS 干扰法及乙基化干扰法提供的信息量还要大，但同样地，缺口的产生可改变 DNA 位点的结构，这种结构改变会进而影响亲和力。

化学置换法：虽然在技术上有难度，但碱基类似物或化学置换却是确定某些置换的化学基团对亲和力影响的好方法。在该例子中，选择性地取代碱基上的某些功能基团，

同时保持碱基的剩余结构完整。利用 K_d 分析测量置换的结果。碱基类似物置换可以尽可能地减少潜在的结构扰乱（见 Lesser et al. 1990）。在一个例子中，利用次黄嘌呤核苷取代鸟嘌呤核苷来研究 TBP 结合（Starr and Hawley 1991）。TATA 盒中的胸腺嘧啶和腺嘌呤分别被胞嘧啶和次黄嘌呤核苷所置换，这些置换使 TATAAAA 的大沟转变为序列 CGCGGGG，但未改变小沟中化学基团的排列（仍与大沟为 TATAAAA 时的小沟化学基团的排列相同）。结果显示，置换对于 TBP 的结合几乎没有影响，从而表明 TBP 主要结合在小沟中。

DNA-蛋白质接触的热力学入门：为了理解通过诱变改变氨基酸-碱基对相互作用的热力学后果，就需要回想一下一个典型的氢键在水介质中提供 1.5 kcal 的能量（带电荷时为 3～5 kcal），而一个范德华键提供约 1 kcal 的能量。根据吉布斯（Gibbs）方程，我们可以计算自由能下降和亲和力下降之间的关系。如果 $\triangle G=RT\ \ln K_d$，则在 25℃ 时，$\triangle G=(1.98)(273+25)\ 2.3\ \log K_d$，其中每个单位都处于标准状态。亲和力每下降 10 倍（即 K_d 增加 10 倍），则伴随着 1.36 kcal 自由能的改变。因此，丧失一个特异性接触（如一个氢键）对结合具有不利影响。诱变方法的一个警告是，改变碱基也会对位点的总体结构产生细微的影响，即不同的弯曲、扭转或转动角都对 DNA 识别或位点形变度起作用（如 *Eco*RI）。当对数据，特别是以弯曲或扭结 DNA 的蛋白质取得的数据进行解释时，一定要牢记这一点。

建立 DNA-蛋白质相互作用的模型

为了说明化学和核酸酶探测技术如何用于研究蛋白质-DNA 相互作用，我们选择了两个表征得很好的例子：GAL4 和 TBP。在这些例子中，可将化学或“足迹表型”的探测（Yang and Carey 1995）直接与蛋白质-DNA 结构比较。GAL4 结合在大沟中，而 TBP 主要与小沟相接触。在这两个例子中，从化学和核酸酶保护分析的结果中推断出 DNA 结合机制的一些重要方面，并随后通过共复合物的晶体结构分析确证了这些推断（Carey et al. 1989；Lee et al. 1991；Starr and Hawley 1991；Marmorstein et al. 1992；Kim et al. 1993a，b）。

GAL4：利用化学和核酸酶保护法对 GAL4-DNA 相互作用进行建模，并从这种建模中做出预测；其中许多预测被 GAL4 DNA 结合结构域及其 17 bp 位点的晶体结构分析所证明。图 11.15 概括了 GAL4 的结构域组织及其图示形式的晶体结构的基本元件。之后，图 11.16 将晶体结构中发现的接触与从化学探针推断的接触进行了比较。

DNA 结合结构域位于前 94 个氨基酸中，可以细分为氨基酸 1～65 之间的 DNA 识别结构域（Marmorstein et al. 1992）和氨基酸 66-94 之间的二聚化结构域（图 11.15A）（Hidalgo et al. 2001；Hong et al. 2008）。氨基酸 11～38 之间的 6 个 Cys（加框的）嵌入在 DNA 识别结构域内，它们螯合两个锌离子，使该区域折叠成能识别 GAL4 结合位点的结构。Cys 和它们的对应蛋白质折叠通常称为双核簇，GAL4 是一个真菌蛋白大家族的创始成员，具有保守的 DNA 识别模体（Schwabe and Rhodes 1997；Todd and Andrianopoulos 1997）。虽然 DNA 识别结构域在溶液中是单体，但在 DNA 上则形成二聚体，这种二聚体的结合能力非常弱，位于氨基酸 65～94 之间的天然二聚化区域能大大增加 GAL4 对其位点的亲和力（Carcy et al . 1989；Hidalgo et al. 2001）。但是，当尝试

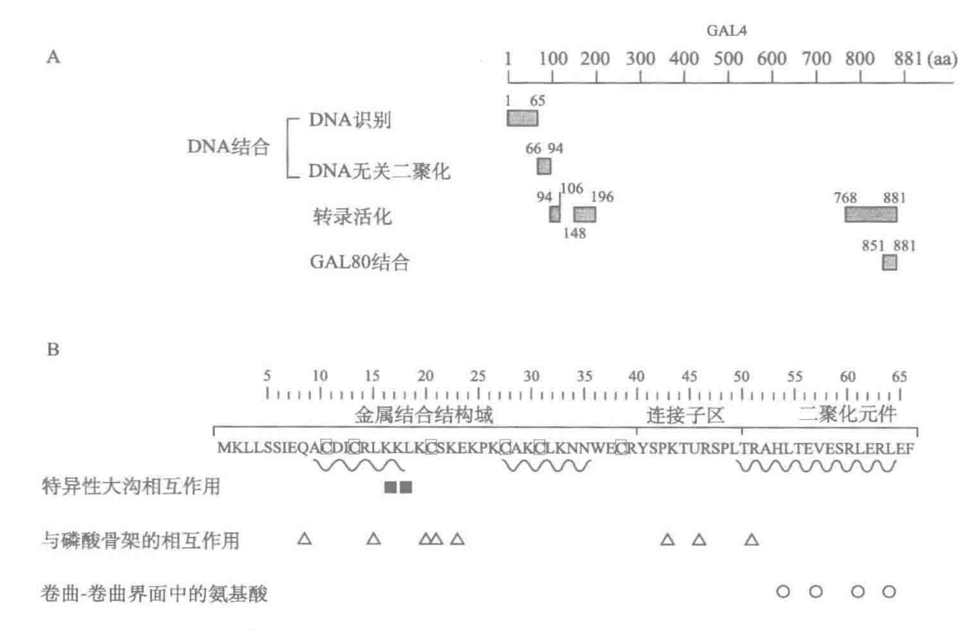


图 11.15　GAL4 及其 DNA 结合位点。(A) GAL4 的结构域组织。(B) 图示形式的晶体结构特征。(B 经允许改编自 Marmorstein et al. 1992，© Macmillan Magazine)

结晶 GAL4 时，含有氨基酸 1～65 的片段和 DNA 形成了晶体，Marmorstein 等（1992）确定了其结构。最新的 GAL4（1～100）结构还揭示了其二聚化结构域的形状（Hong et al. 2008）。

图 11.16A 概括了 GAL4 在 DNA 上的化学和核酸酶足迹结果（Carey et al. 1989）。二分体位点与 GAL4 接触的对称性最初表明 GAL4 以二聚体形式与其 17 bp 位点结合，其中每个单体与二分体位点的一半接触。但是，当从 DMS、羟自由基和乙基化干扰法中推断的相互作用展示在典型的 B-DNA 螺旋上时（图 11.16B，左图），这些相互作用就揭示出重要信息，这些信息导致了更精密模型的建立，该模型说明了 GAL4 是如何结合其位点的。大多数大沟 DMS 保护或干扰发生在位于二分体位点两端的 G 残基上（灰色的 G 残基），而小沟或骨架相互作用从这些 G 向位点的中央延伸。被乙基化干扰牵扯到的受影响磷酸沿着骨架表示为黑色，被羟自由基足迹牵扯到的受影响糖残基表示为粗体棍棒。根据异源二聚体数据（图 11.4）和化学数据（图 11.16A），Carey 等（1989）提出 GAL4 的每个单体“与识别位点较靠外的碱基对在大沟中发生序列特异性接触，然后沿磷酸骨架的一条链蛇行”，其中，“二聚体的接触位于二分体位点的中央”。值得注意的是，已经解析的结构（Marmorstein et al. 1992）确证并显著完善了这一简单模型。该结构表明由化学探测法鉴定的相互作用和晶体结构中鉴定的相互作用非常吻合。

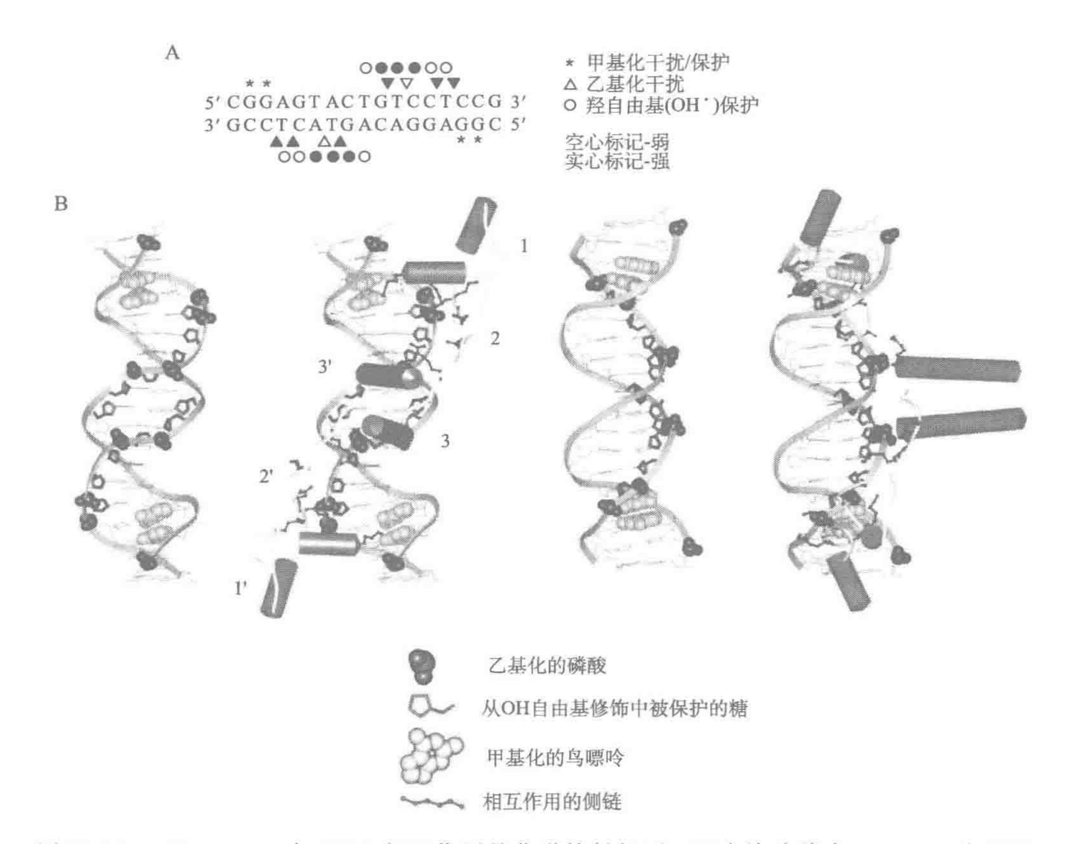


图 11.16 (A) GAL4 与 DNA 相互作用的化学接触概要（经允许改编自 Carey et al. 1989, © Academic Press)。(B) 相互作用的晶体结构。显示了两个旋转视图（由 Michael Haykinson [UCLA] 利用 Molecular Graphics structure modeling computer program Insight II 绘制）。

通过将 GAL4 的晶体结构叠加到 DNA 上，并且标示出从化学探测法推断的接触，最好地说明了这一点（图 11.16B，右图）。首先，很明显，结构中的 GAL4 直接与 DNA 一些残基相互作用，而这些残基都涉及化学探测发现的接触。更具体地说，蛋白质的离散结构特性产生与 DNA 的接触。锌复合的 DNA 识别模块位于氨基酸 8～40 之间（图 11.16B)。在这个结构域中，每个单体中的 2 个锌原子与 6 个 Cys 复合，折叠成含有 2 个互相垂直的 α 螺旋结构（图 11.16B，特征 1 和 1′)。来自每个单体的一个 α 螺旋插入到垂直于 DNA 轴的大沟中，在 GAL4 的 Lys17 和 Lys18 之间（表示为从圆柱形螺旋伸出的黑色棍棒）及二分体位点末端 G 残基的暴露基团（表示为灰色的范德华键）发生序列特异性相互作用。这些 G 就是在化学探测实验中通过 DMS 保护和干扰分析检测到的 G（图 11.16A)。在真菌双核簇转录因子超家族中，多数成员都在第 17 位含有保守的 Lys 或 Arg，在第 18 位含有保守的 Lys 或 His。对于很多已知蛋白质来说，这些残基和二分体 CGG 模体的接触是 DNA 识别机制的保守特征（Liang et al. 1996，Schwabe and Rhodes 1997，Swaminathan et al. 1997)。因此，GAL4-DNA 化学探测数据和结构可应用于这一非常大的因子家族，从而有助于确定这些因子如何识别其位点。

每个单体中紧挨 DNA 结合模块的羧基末端是一个易变的、非结构化的连接区（2 和 2′），该连接区沿小沟排列。在该连接区中，主要是 Lys 和 Arg 参与同磷酸骨架的非特异性接触，直至到达二分体位点的中心（图 11.16B）。黑色棍棒表示突出到 DNA 上的侧链，这些侧链与乙基化干扰和羟自由基保护牵扯的残基相互作用（磷酸显示为黑色的范德华键，糖以棍棒显示）。在双核簇家族成员中该蛋白质连接区的长度不同。该长度的变化使不同成员与 5′ 和 3′ CGG 模体之间距离不同的位点相结合（Schwabe and Rhodes 1997）。

两个单体都在二分体位点的中央形成一个 α 螺旋（3 和 3′）。利用形成疏水界面的 Leu 和 Val，来自每个单体的 α 螺旋相缔合，形成一个卷曲螺旋的二聚化模体（图 11.16B）。这个卷曲螺旋模体基部参与与二分体位点中央小沟的非特异性接触。同样，由乙基化干扰及羟自由基干扰表明的接触与晶体结构中卷曲螺旋的底部相互作用的残基相匹配。综上所述，这些研究结果表明从化学探测数据预测的模型（Carey et al. 1989）与晶体结构非常吻合（Marmorstein et al. 1992）。

TBP：TATA 盒结合蛋白（TBP）是只在小沟中结合并与碱基对和磷酸骨架发生相互作用的蛋白质的实例。不同物种间 DNA 识别序列的比较表明，长度为 180 个氨基酸的相对保守性 TBP 核心序列，对于 TATA 盒与通用转录因子的结合及相互作用是必需的。这段核心序列包含了两个 80 个氨基酸的直接重复。TBP 基因的氨基端在种间差异很大，据推断，该区域介导 TBP 与转录机器组分的物种特异性相互作用，而转录机器组分是转录激活所必需的（Burley and Roeder 1996）。

在解析与 TATA 框相结合的 TBP 晶体结构之前，研究者进行了大量的化学足迹分析，以理解 TBP 和 TFIID 如何与 TATA 盒对接。这些分析导致一种模型，该模型的主要元件被后来的晶体结构所证实。图 11.17A 以图示的方式说明了羟自由基足迹、乙基化干扰及 DMS 保护和干扰分析的结果。请注意，从化学探针推断而来的大沟相互作用（即 N7-甲基化 G 残基）的缺乏及假定的小沟接触（即 N3-甲基化 A 残基和乙基化磷酸）的优势。对这些数据的最初解释作为 TBP 主要和小沟接触这一模型的依据（Lee et al. 1991；Starr and Hawley 1991）。

有两个实验室分别独立地解析了 TBP/TATA 共复合物的晶体结构，这些晶体结构证实并扩展了这一模型，如 Sigler 及其同事解析了酵母 TBP 共晶体（Kim et al. 1993b），Burley 及其同事解析了拟南芥（*Arabidopsis*）TBP（Kim et al . 1993a）。拟南芥 TBP 的结构表明，TBP 结合 TATA 的机制在整个进化过程中是保守的（Patikoglou et al. 1999）。在图 11.17B 中，左边的图显示了将化学接触叠加到了 DNA 上（两个旋转视图），而右边的图显示了带有 DNA 的 TBP 结构。该结构含有 10 条反向平行的 β 折叠片（箭头），呈现马鞍状构象，其中 4 个 α 螺旋（圆柱）冠盖在马鞍的上表面。该蛋白质由两个亚结构域构成（图 11.17B，1 和 1′），分别对应于一级结构中的每个直接重复。每个亚结构域组织成“折叠片 1-螺旋 1 -折叠片 2 -转角-折叠片 3 -折叠片 4 -折叠片 5 -螺旋 2”的构型（图 10.7；见第 10 章）。组成马鞍下表面或凹陷面的 β-链形成一个大的疏水界面，该疏水界面通过范德华键和疏水性接触与碱基相互作用。在小沟中的 8 bp 区形成数量很少的特异性氢键（Sigler 及其同事鉴定有 6 个，Burley 及其同事鉴定有 4 个）。最近的研究对这些最初的发

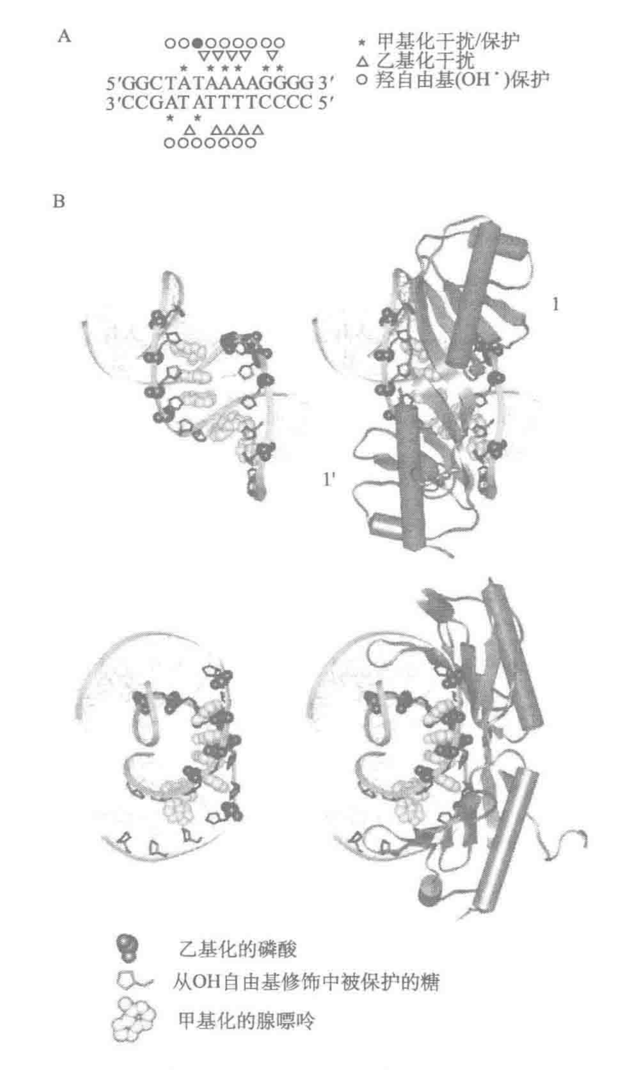


图 11.17 （A）TBP-DNA 相互作用的化学接触总结。（B）晶体结构（由 Michael Haykinson［UCLA］利用 Molecular Graphics structure modeling computer program Insight II 绘制）。

现做了更深入的阐述（Patikoglou et al. 1999）。结合引起 DNA 的显著变形，导致 80°弯曲，使 DNA 解旋 110°，导致图中所示的特征性弯曲形状。这种变形是由两个 Phe 插入到 TATAAA 两端的碱基之间引起的。这种插入作用使碱基弯曲，从而导致整个位点弯曲。仔细观察由灰色范德华键表示的 DNA 中的两个上部的 A 残基，就可以想象到这一点。带有停泊的 TBP 的图片显示了插入到两个 A 之间的插入性 Phe，Phe 的插入导致两个 A 的弯曲。产生的弯曲和解扭使蛋白质和小沟间的疏水界面达到最大。碱性残基和磷酸之间的盐桥、水分子介导的氢键，以及与糖的范德华接触都是相互作用背后的驱动力。最重要的是，图 11.17B 表明当该结构与 DNA 对接时，晶体相互作用解释了羟自由基保护和干扰

TBP 结合的几乎所有的乙基磷酸。

上面讨论的 GAL4 和 TBP 的实验是两个研究实例，在这两个实例中，化学探测实验的细节使研究者提出蛋白质与 DNA 如何相互作用的模型。一旦结构被解析，这些结构就可以证实并扩展原始模型的基本环节。更多的结构被解析后，DNA 结合的化学和机制之间的相互关系应能提供足够的信息，从而使研究者对数据进行推算，以理解未知蛋白质或属于已界定家族的蛋白质是如何介导 DNA 结合的。

启动子特异性多组分核蛋白复合物的分析

本章开始时，我们描述了与启动子及增强子（即增强体）结合的蛋白质簇如何被最好地被描述为蛋白质-蛋白质和蛋白质-DNA 相互作用的网络（见第 1 章和 Carey 1998）。增强体复合物形成背后的两个驱动力是协同性和 DNA 成环或弯曲。协同性是 DNA 结合中蛋白质互相帮助的一种现象，以便蛋白质复合体比单独的每种蛋白质都结合得更为紧密。换句话说，有其相互作用的搭档存在时，一种蛋白质表现出对 DNA 位点的更低的 K_d 值或更高的亲和力。DNA 成环是当两个远距离蛋白质相互作用时导致的 DNA 的弯曲。在 DNA 的构象保持长度（约 140 bp）内，DNA 成环榨出的能量损失必须由蛋白质-蛋白质相互作用力付出，或者由一个蛋白质付出，该蛋白质通过弯曲两个相互作用的蛋白质间的 DNA 而稳定 DNA 环。DNA 弯曲被定义为 DNA 曲度的局部改变，这种改变由 DNA 序列本身或通过一种改变 DNA 螺旋轨迹的蛋白质引起。一旦两个相互作用的蛋白质之间的距离超过了 DNA 的构象保持长度，能量损失就不会太可观，但蛋白质-蛋白质相互作用的可能性会随距离的增加而减小。生物物理化学家已经建立了这些效应的模型，但才刚刚开始以预测的方式应用于转录系统（Wang and Giaever 1988；Rippe et al. 1995）。

DNA 结合的协同性

形成特异性的关键机制之一是协同性 DNA 结合。鉴于增强体，这一概念在真核基因表达领域正在被重新审视。在许多 DNA 结合蛋白协同性结合的例子中，具有代表性的例子是 λ 抑制因子或 *c*I（Johnson et al. 1981）。利用常规生物化学和生物物理技术已经分析了 λ 抑制因子的结合，并对结合平衡进行了数学分析，为理解和研究协同性提供了良好的模型（综述见 Hochschild 1991；Senear et al. 1998）。在噬菌体溶源中，抑制因子通过控制从两个相关的向右和向左操纵子（分别为 O_R 和 O_L）的转录，维持溶源状态。为简单起见，我们描述一下在 O_R 中的情况。在 O_R 情况下，λ 抑制因子二聚体协同性地结合高亲和位点（O_R1）和低亲和位点（O_R2）。协同性由相邻二聚体的羧基端结构域之间的直接蛋白质-蛋白质相互作用介导（综述见 Koudelka 2000）。结合在 O_R1 处的抑制因子抑制激活溶菌周期的右向启动子（P_R）的转录。同时，结合在 O_R2 处的抑制因子与其靶点（即 RNA 聚合酶）相接触，并刺激一个维持抑制因子的启动子（P_{RM}）的转录。P_{RM}控制 λ 抑制因子的表达，是一个经典的自调控启动子的例子。

协同效应最初是由DNase I足迹分析测定的。通过比较抑制因子对[32]P末端标记的、在其天然位置含有O_R1和O_R2或单独的O_R1或O_R2的DNA片段的亲和力，Ptashne及其同事发现抑制因子协同性地结合（Ptashne 1992）。如前文所述，DNA片段的浓度设置在K_d以下。然后以两倍阶梯滴定抑制因子，直至DNA片段被饱和，就如通过DNase I足迹分析测量一样。回想一下，如果DNA浓度在K_d以下，那么当[PS] = [S]（见专题11.1），即DNA位点被50%占据的条件，则K_d= [P]。简单来说，K_d等于产生50%占据位点所需的游离蛋白质的浓度。利用这种方法，表明λ抑制因子对于单独的O_R1和O_R2的K_d值分别是3.3×10^{-9} mol/L和5×10^{-8} mol/L（图11.18A）。但是在O_R1存在时，对O_R2的K_d降低至大约10^{-9} mol/L。协同性的程度可以表示为存在和不存在O_R1的情况下测量的O_R2 K_d的比值。因此，在评价蛋白质是否协同性结合时，必须测定它们单独和一起时的亲和力。在不同的蛋白质协同性结合的情况下，可以利用相同的DNA模板进行该分析。然而，在一种蛋白质与多个位点结合的情况下，必须将蛋白质对野生型模板的结合与对一个位点或另一个位点被突变的模板的结合相比较。一般来说，协同性很容易通过DNase I足迹的EMSA观察到。在这种情况下，一种蛋白质会以大于加和性的方式增加被另一种蛋白质所迁移的探针量。通常来讲，使用结合DNA和未结合DNA的Hill图，可定量这些数据。在其他较旧的详尽的综述中包含这一现象的量化模型（Brenowitz et al. 1986；Hochschild 1991；Senear et al. 1998）。

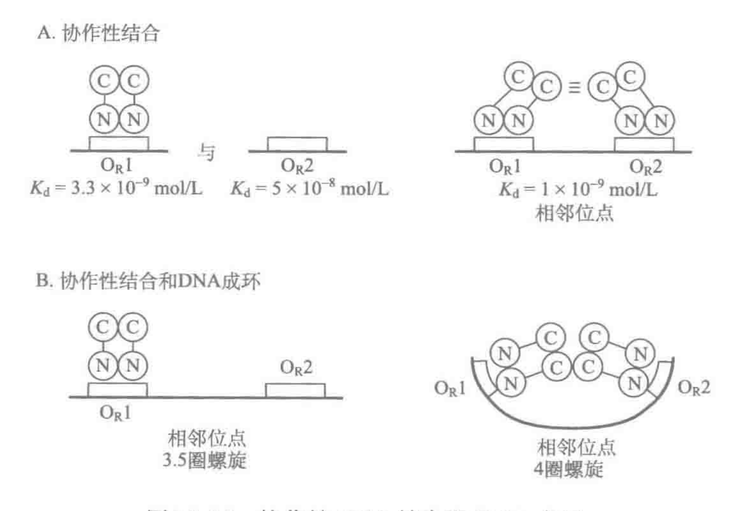


图11.18 协作性DNA结合和DNA成环。

在一个多组分复合物中，理论上所有的蛋白质-蛋白质相互作用都应对彼此的稳定性具有互惠效应。在某些情况下，这些效应是明显的（如与亲和力弱的位点结合的两种蛋白质），而在其他情况下，这些效应则不明显（如结合在O_R2上的λ抑制因子对结合O_R1具有三倍效应）（Johnson et al. 1981）。在考虑增强体可能如何工作时，可以预测很多相互作用将是协作性的。作为这一思路的延伸，可以想象当启动子结合的激活因子促进通用转录机器组装成转录复合物时，该复合物应转而辅助激活因子与DNA上它们

的位点相结合。IFN-β 增强体及 Epstein-Barr 激活因子 ZEBRA 似乎就是这种情况（Kim and Maniatis 1997；Eliwood et al. 1999）。关于多蛋白复合物结构的综述揭示了涉及协同性增强体组装的很多原理（综述见 Garvie and Wolberger 2001；Ogata 2003）。

DNA 成环和弯曲

DNA 成环：当两种蛋白质被结合到 DNA 上时，它们相互作用的能力就形成了 DNA 成环的理论基础，长期以来，一直认为 DNA 成环是远端结合的激活因子与通用转录机器相互作用的机制。DNA 成环的概念是在对 *E. coli* 中阿拉伯糖和乳糖操纵子的研究中首次提出的，并且已经涉及大量的生物过程。已利用 λ 抑制因子系统对该现象进行了广泛的研究，下面我们就以此为例进行讨论。尽管用这些原核系统充当增强子作用的非常有趣的模型，但它们不能完全解释真核细胞中在很远的距离上发生的相互作用的类型。正如第 1 章中所讨论的，涉及染色质成环和跟踪机制的非常复杂巧妙的机制，将远端的增强子和启动子带到一起（综述，见 Alvarez et al 2003；Bondarenko et al. 2003；Dekker 2003）。因此，下面描述的相互作用主要是为了便于分析位于 150～200 bp 内的蛋白质之间的相互作用。

如以前所讨论的，正常情况下，λ 抑制因子与两个中心距离（包括这个距离）为 25 bp 或 2.4 个螺旋的 17 bp 位点协同性结合。插入 10 bp 或 11 bp（称为螺旋增长，因为 DNA 的周期是每圈为 10.5 bp）对 8 圈 DNA 螺旋以外的协同性没有影响。但是，在距离 8 圈螺旋的位点之间插入非螺旋性片段，则破坏了协同性。图 11.18B 图示说明了一个例子，该例子中在抑制因子结合位点之间插入 3.5 圈螺旋阻止了协同性，而插入 4 圈螺旋则使协同性得以恢复（Hochschild and Ptashne 1986）。电镜和物理学测量已经证实了两位点间的 DNA 确实成环（Griffith et al. 1986）。平面内的成环仅需要两位点间 DNA 的曲度发生平稳的改变，以便相邻的蛋白质能够接触。DNA 曲度的值取决于蛋白质之间的距离、相互作用蛋白质的大小，以及它们相互作用的位置。非螺旋的引入导致的蛋白质不能相互作用，是由扭转 DNA 的能量损失造成的。在上文所述的距离上，需要消耗 7 kcal 或更多的能量。据推测，相邻二聚体间的蛋白质-蛋白质相互作用的能量可能不足以维持相互作用及承担扭曲所需的能量消耗。这种插入受到一种观察现象的支持，即在位点间引入一短的单链缺口（原则上这一缺口要允许围绕该单链上的自由旋转），即使插入一段非螺旋增长也能恢复协同性。令人惊异的是，多达 20 圈螺旋也能观察到协同性，但对螺旋周期性的需求只发生至 8 圈。据推测，8 圈之后，扭转的能量损失较小，并且相互作用不再受到限制（综述见 Rippe et al. 1995；Travers 2004）。在分析认为有两种或多种蛋白质相互作用的情况时，改变螺旋相位的效应一般可用于某种相互作用的证据。

虽然从发生在远距离并导致一个蛋白质与 DNA 协同结合的蛋白质-蛋白质相互作用可以推断 DNA 成环，但是难以直观地显示出 DNA 成环。然而，结合位点间 DNA 的 DNase I 切割模式却能提供这一方面的某些信息。在抑制因子结合被 6 圈螺旋隔开的位点的情况下，位点间的 DNA 表现出具有 10 bp 周期性的一系列增加和保护（Hochschild and Ptashne 1986）。对这一现象的解释是 DNA 被平滑地弯曲，从而在弯曲的外凸面上暴露出小沟，并保护了内凹面上的小沟。这种切割模式可能就是成环作用的判断特

征。增强体中激活因子之间的远程相互作用可能也表现出这种周期性。

λ抑制因子模型系统有几个特征可能与真核系统不同。有些蛋白质可能大而足够柔韧，以致无需能量损耗就能相互作用，并与它们的螺旋关系无关。另外，可以想象蛋白质-蛋白质相互作用的强度足以承受能量消耗。在其蛋白质-蛋白质相互作用的意义上，λ抑制因子确实代表一种新的蛋白质，虽然它不能承受将蛋白质带入相位的DNA扭转的能量消耗，但它能承受同相DNA弯曲和成环的能量消耗。有很多情况并非如此的相互作用，包括TCR-α和IFN-β增强体中激活因子之间的相互作用，这些相互作用需要建筑蛋白，而这些建筑蛋白可以显示恰当塑造DNA的基本序列特异性（Grosschedl 1995；Carey 1998）。

原子力显微镜是另一种研究DNA成环的方法，该方法已经被应用于研究复合物中的蛋白质-蛋白质相互作用。一个特别好的例子是在造血细胞中表达的c-Myb和C/EBP-β在*mim*-1启动子上的相互作用。Ogata及其同事将显示了相互作用表面的c-Myb-C/EBP-β复合物结晶学与原子力显微镜相结合，鉴定了在溶液中两个蛋白质之间形成的85 bp的DNA环（Tahirov et al. 2002；Ogata et al. 2003）。

通过建筑蛋白的DNA弯曲机制

在原核和真核系统中的研究已经阐明了DNA弯曲在DNA代谢过程中所发挥的重要作用。在转录过程中，DNA弯曲将紧密就位的蛋白质排列成与染色质重塑因子及通用聚合酶II转录机器接触的复合物。尽管不同阶梯的碱基具有不同的弯曲能力，但在DNA的构象保持长度（140～150 bp）内，被动的DNA弯曲由于DNA螺旋的刚性而受到限制（综述见Travers 2004）。虽然某些蛋白质-蛋白质相互作用可通过位点间DNA的被动成环而发生，但这些相互作用需要由该蛋白质-蛋白质相互作用给予的阈值量的能量。如果该能量不充足，蛋白质只能在第三种蛋白质的帮助下发生相互作用，其中，第三种蛋白质具有将DNA按某一方向弯曲到特定角度从而使该相互作用发生的能力。这些弯曲蛋白即被称为建筑蛋白（Werner and Burley 1997）。

在真核细胞中，DNA弯曲蛋白特异性地或非特异性地结合DNA（Werner and Burley 1997）。特异性蛋白质的例子包括TBP、LEF-1（参与TCR-α增强体形成）、性别决定因子SRY及参与IFN-β增强体形成的HMG（I）Y。LEF-1含有称为HMG B-box的结构域。该结构域首先在真核HMGB1和HMGB2蛋白中被鉴定，这两种蛋白质是非特异性DNA弯曲蛋白的典型例子。HMG B-box蛋白与小沟相结合。HMG B-box的晶体结构和NMR结构显示了具有三个α螺旋的L形结构（图11.19右）。在碱基对之间插入一个疏水性的氨基酸会促进弯曲，从而造成碱基的解堆积、螺旋的部分解旋，以及向大沟的弯曲。虽然LEF-1是一个序列特异性HMGB蛋白的例子，但在酵母和人类中某些最丰富的蛋白质都是非特异性的，包括来自哺乳动物的HMGB1和来自酵母的NHP6A和NHP6B。NHP6A与小沟的结合与LEF-1非常相似，并有一个环围绕DNA到达大沟（图11.19左）。

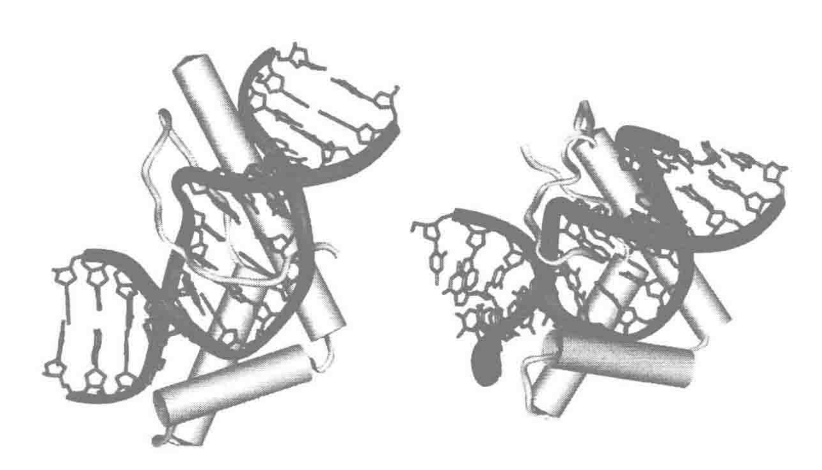


图 11.19　结合在小沟中的序列特异性 LEF-1（右）及引起的 DNA 弯曲，并与 NHP6A（左）的非特异性结合相比较［由 John Heiss（UCLA）用 Molecular Graphics structure modeling program Insight II 绘制］。

并非所有弯曲 DNA 的蛋白质都结合在小沟，最典型的例子是原核调控蛋白 CAP 和 Fis（Pan et al. 1996）。这些二聚化蛋白质通过将识别 α 螺旋插入邻近的大沟中，使 DNA 发生弯曲。在 Fis 的情况下，识别螺旋中心间的距离小于大沟中心间的距离，因此蛋白质必须弯曲 DNA，以便 DNA 位点与蛋白质接触。在这种情况下，DNA 也可能更广泛地包裹在蛋白质周围，进一步扩大弯曲。有些真核蛋白质，包括 bZIP 家族成员，也会使 DNA 从大沟发生弯曲（Hardwidge et al. 2003）。在这种情况下，该机制似乎与蛋白质的碱性侧链对磷酸的中和有关，从而遮蔽不利的磷酸盐斥力。这反过来会促使 DNA 倒向蛋白质。

研究 DNA 弯曲的方法

测量 DNA 弯曲的三种常用方法是 EMSA、DNA 环化（综述，见 Crothers et al. 1991，1992；Kahn and Crothers 1992）和 FRET（Hillisch et al. 2001）。EMSA 是最容易处理且信息量最大的方法之一，因为与已知的标准相比，它可以提供关于弯曲的位置、程度和方向的信息。

当 DNA 在聚丙烯酰胺或琼脂糖凝胶中迁移时，据认为 DNA 要蜿蜒滑行通过凝胶小孔。虽然不同长度的分子具有同样的电荷密度（归因于糖/磷酸骨架的均一性），并因此经受相同的电场力，但较小的分子能有效地通过凝胶孔。弯曲的 DNA 迁移率变慢，同时当弯曲定位于 DNA 分子的中央时，相对于弯曲定位于 DNA 分子的两端，这种迁移率变慢的效应更加显著。目前的观点是迁移率与 DNA 两端间距离的平方成反比。在中央部位发生弯曲的 DNA 分子与线性 DNA 分子或靠近末端发生弯曲的 DNA 分子相比，两个末端间的距离最小，因此迁移率的滞缓程度最大。

在一个典型的分析中（图 11.20），将目的 DNA 位点克隆到含有两个被一独特克隆位点隔开的串联多聚接头的载体中（Zwieb and Adhya 1994 ）。然后，用限制性内切核酸酶切割 DNA 分子，其中内切核酸酶在每个多聚接头上只切割一次。这样就产生了一系列被循环改序的片段，其序列组成相同，但位点相对于多聚接头末端的位置不同。然后将 DNA 分子用^{32}P 标记，与蛋白质一起孵育，继而使复合物在聚丙烯酰胺凝胶中分级分离。另一种方法是用 PCR 生成改序的片段。相对于未结合的 DNA，蛋白质滞缓了所有的片段。然而，如果蛋白质使 DNA 螺旋明显弯曲，那么它对含有中央定位位点的片段滞缓的程度将大于含有末梢定位位点的片段。片段的迁移率（垂直轴）测量为片段到凝胶末端的距离，然后将迁移率相对于 DNA 片段中位点的位置进行作图。曲线会显示出中央弯曲的片段迁移率最小。

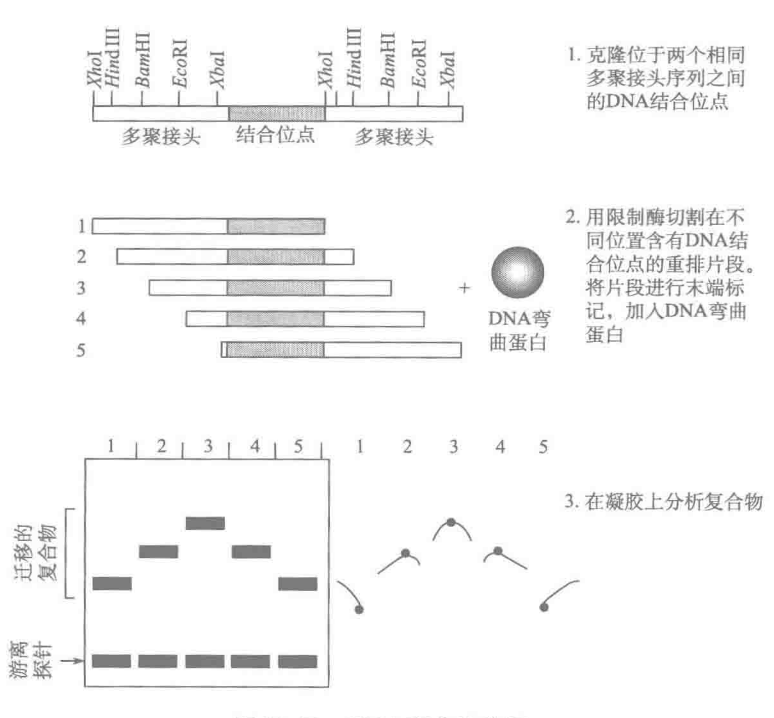


图 11.20 DNA 弯曲的分析。

为了确定弯曲的方向性，将结合位点置于含有定相的 A 区（A tract）的 DNA 片段上。定相的 A 区是 5 个或更多个连续的 A 残基的延伸，这些 A 残基导致 DNA 向已知方向弯曲。然后将不同长度的接头置于 A 区和蛋白质结合位点之间。随着接头长度的延长或缩短，蛋白质引起的弯曲和 A 区弯曲彼此相对转动。在这种情况下，片段会表现出正弦曲线的迁移率模式。当弯曲同向时，片段的迁移率达到最小。这时，可以确信蛋白质引起的弯曲与 A 区引起的已知弯曲方向一致。可以在天然位点的前后序列中评价这种弯曲的生物相关性。

A 区弯曲的角度已经严格地确立，并且因此可以作为一个标准，以确定由蛋白质引起的弯曲。在该实验中，将蛋白质结合位点再一次置于含有定相的 A 区的 DNA 分子一端。随着 A 区的数量增加，弯曲的角度也增加。因为 A 区都由 10～11 bp 隔开，每个 A 区的弯曲方向都是相同的，那么随着 A 区的增加，弯曲也增大。末端中蛋白质结合位点的存在允许对迁移率进行校正。现在，将带有中央部位的弯曲位点的蛋白质-DNA 复合物与 A 区标准进行比较（Crothers et al. 1991）。当使用 EMSA 时，要注意有些蛋白质在凝胶迁移所必需的条件下不能存活，并且含有大蛋白质的 DNA 复合物可能难以在聚丙烯酰胺凝胶上解析。在这种情况下，基于液相的 DNA 环化分析可能是有用的替代分析方法。

DNA 环化是研究弯曲的另一种经典方法。通常将蛋白质结合位点设计位于一条 150 bp 的 DNA 片段的中心位置，该片段是通过产生黏性末端的限制内切核酸酶剪切一个 DNA 分子产生的。由于 DNA 在构象保持长度（persistence length）内的刚性，150 bp 的分子以低频率环化。加入使 DNA 分子弯曲的蛋白质使 DNA 分子的两个末端靠近，从而有利于通过 DNA 连接酶将 DNA 分子的两端连接起来。然后除去蛋白质，DNA 环化产物在 5%～8%聚丙烯酰胺凝胶中进行电泳分析。电泳时，环化产物在凝胶中的迁移速度比线性的 DNA 慢很多，并且可以在蛋白质存在和不存在时生成的环化产物的量可以用来评价 DNA 弯曲。此外，带有取代蛋白质结合位点的 A 区的类似的 DNA 分子在平行反应中也被连接，并且可以将其环状产物的积累速率与含有蛋白质的反应的速率相比较。A 区也可以用来确定弯曲的方向，因为 A 区与蛋白质引起的弯曲同相，从而导致 DNA 两末端更加靠近，提高连接效率（综述，见 Crothers et al. 1992；Kahn and Crothers 1992）。

FRET 也可以用来研究 DNA 弯曲。与 EMSA 和 DNA 环化法不同，FRET 是一个非常复杂的方法，有很多注意事项。它通常用来确认其他方法的研究结果，或当 EMSA 方法失败但怀疑 DNA 弯曲时使用。由于 FRET 的距离限制，通常使用的 DNA 序列显著地短于环化和 EMSA 实验中使用的 DNA 序列。标记在 DNA 分子两端的染料分子（即 Cy3、荧光素和四甲基罗丹明）的位置也是已知的（Hillisch et al. 2001），因此，特殊的染料对可以标记在 DNA 的两端。当蛋白质使 DNA 弯曲时，DNA 两端便更加接近，并发生 FRET。有两种方法来测量 FRET，包括测量来自受体荧光基团的发射或测量受体对供体的淬灭。FRET 已经被用于测量 TBP（Parkhurst et al. 1996）和 HMGB 家族成员（Lorenz et al. 1999）引起的 DNA 弯曲。

总之，对真核激活因子如何与 DNA 结合并组装成增强体的协调一致的研究，需要投入很大的精力将多种方法应用于高度繁杂的问题。在原核领域中，这些复合物的研究花了 10 多年的时间，所需要的投入是巨大的。然而，这种核蛋白（nucleoprotein）复合物的组装和调控几乎可以肯定地说，它们掌握着组合控制和基因调控过程中特异性现象的钥匙。

技术

方案11.1 DNase I足迹

DNase I足迹方案是1978年提出的（Galas and Schmitz 1978），该方法以嵌套理论为基础，该理论也是DNA测序的概念基础（Maxam and Gilbert 1977）。由于其简单性，DNase I足迹法已被广泛用于鉴定和表征DNA-蛋白质相互作用。该方法的概念是一个^{32}P单末端标记的片段的DNase I部分消化，会产生梯状片段，这些片段在变性聚丙烯酰胺凝胶上的迁移率及其在随后的放射自显影中的位置，将代表从末端标记到切割位点的距离。结合的蛋白质阻止其结合位点内及位点周围DNase I的结合，从而形成切割的梯状条带中的“足迹”（专题11.3和图11.9）。从末端标记到足迹边缘的距离表示DNA片段上蛋白质结合位点的位置。位点的确切位置可以通过将DNA测序梯带和足迹分析样品一起电泳确定。

由于空间位阻，DNase I不能直接结合在DNA结合蛋白的邻近部位，因此，足迹表示的结合位点较宽，一般比位点本身大8～10 bp。而且对DNase I的晶体结构和广泛的生化研究（综述，见Suck 1994）表明它结合在小沟中，与糖/磷酸骨架的两条链相接触，并使DNA向大沟弯曲。特定位置上小沟的序列依赖性宽度会影响切割效率。因此，切割梯带并不是均匀的，会存在某些脱节。有些情况下，蛋白质会结合在这些脱节的片段上，这些位点上的足迹不如结合位点位于切割较均匀且有效的区域上时明显。足迹分析可以利用纯的蛋白质进行，也可以利用提取物进行。下文及第7章中详细说明了对纯的和粗的系统的技术考虑。

时线与安排

一般来讲，必须先制备含有蛋白质结合位点的单末端标记的DNA片段。对于小的位点来说，一条单链被标记的双链寡核苷酸（30～35 bp）就符合要求。还必须有高活性的DNase I储液，DNase I储液可以由DNase I粉末或溶液制备。整个分析通常花费三天，在利用目的蛋白绘制剂量应答曲线之前，必须首先确定DNase I的最佳浓度。理想情况下，研究者应具有纯蛋白或表达该蛋白质的提取物，以及表明蛋白质确实与位点相结合的来自诱变研究、转染分析或EMSA实验的初步数据。

第1天：制备^{32}P末端标记的片段（见方案11.6）。

第2天：滴定DNase I，为足迹分析优化酶的用量和实验条件。

第3天：用最佳用量的DNase I绘制目的蛋白的剂量应答曲线。

提纲

步骤1：从预制丙烯酰胺/尿素混合液倒胶（30 min）。

步骤2：准备DNase I足迹分析缓冲液（1 h）。

步骤3：准备反应和滴定（30 min），将反应混合液于适当温度下孵育（15 min）。

步骤 4：稀释 DNase I（15 min），加入 DNase I，进行足迹反应（5 min）。终止反应，并对产物进行加工（1 h）。

步骤 5：凝胶电泳（4～6 h）和放射自显影（8～12 h）或磷屏成像分析。

步骤 6：进行 DNase I 滴定数据分析。

步骤 7：绘制剂量应答曲线（1 天）。

步骤 8：进行剂量应答数据分析。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

β-巯基乙醇（14 mol/L）〈!〉

牛血清白蛋白（BSA；无核酸酶，50 μg/μl）

缓冲液 D

　　20 mmol/L HEPES-KOH（pH 7.9）〈!〉

　　0.2 mmol/L EDTA

　　20% 甘油

　　0.1 mol/L KCl〈!〉

　　缓冲液 D（Dignam et al. 1983）是一种非常有用的 DNA 结合缓冲液。核提取物应在含有苯甲基磺酰氟（PMSF）和二硫苏糖醇（DTT）的缓冲液 D 中透析。然而，任何相似组分的缓冲液都与该分析兼容，盐及镁离子的终浓度必须与蛋白质体外结合及 DNase I 的催化活性兼容。

$CaCl_2$（1 mol/L）〈!〉

氯仿〈!〉

DNA 结合蛋白（重组；250 ng/ml～1 mg/ml）

DNA 模板（^{32}P 末端标记；50 fmol/μl，10 000～100 000 cpm/μl）〈!〉

关于末端标记 DNA，见方案 11.6。

DNase I（无核酸酶，FPLC 纯化；5 000～10 000U/ml）

干冰（可选）〈!〉

乙醇（80%和 95%）〈!〉

甲酰胺染色混合液

　　98%去离子甲酰胺〈!〉

　　10 mmol/L EDTA（pH8）

　　0.025% 溴酚蓝〈!〉

　　0.025% 二甲苯青 FF〈!〉

$MgCl_2$（0.1 mol/L）〈!〉

苯酚〈!〉

聚丙烯酰胺/尿素凝胶混合液〈!〉

关于丙烯酰胺凝胶混合及灌注的说明，见 Sambrook 和 Russell（2001）。

Poly（dI-dC）（1 μg/μl）

蛋白质稀释缓冲液

缓冲液 D　500μl

50 μg/μl BSA　1 μl

14 mol/L β-巯基乙醇　1 μl

BSA 在这里用作为蛋白质稳定剂，β-巯基乙醇用作为还原剂；也可以使用 0.1～1 mmol/L 的 DTT。还原剂对于 DNA 结合结构域（即锌指或许多 bZIP 蛋白的碱性区）中具有半胱氨酸的蛋白质特别重要。抗凝剂（如 0.01%～0.1% NP-40 或 Triton X-100）也可证明是有益的。这种缓冲液少量配制，并分装成小份冷冻保存。不推荐在冰上或冰箱里长期保存该溶液。

蛋白酶 K（10 mg/ml）〈!〉

终止缓冲液

400 mmol/L 乙酸钠

0.2% SDS〈!〉

10 mmol/L EDTA

50 μg/ml 酵母 tRNA

该终止缓冲液终止 DNase I 反应，同时准备该缓冲液用于酚抽提和乙醇沉淀。加入 EDTA 是为了螯合 DNase I 活性所必需的二价阳离子；SDS 变性蛋白质，并使其从 DNA 上脱落；乙酸钠和载体 tRNA 促进乙醇沉淀。终止缓冲液可制备成浓缩储液室温下保存数月。

设备

电泳仪

Eppendorf 管（非硅化的 1.5 ml 和硅化的 0.5 ml）

－80℃冰箱（可选）

盖革计数器

凝胶干燥机，真空

微量离心机

巴斯德吸管（窄口）

磷屏成像仪暗盒（可选）

计时器

水浴，设在 30℃

X 光片（可选）

程序

步骤 1：从预制聚丙烯酰胺/尿素混合液倒胶

在实验开始前约 2 h，灌制 8%～12%聚丙烯酰胺/尿素凝胶（根据片段大小确定凝

胶浓度；见 Sambrook and Russell 2001）。凝胶的聚合过程需要 1 h，上样前，预电泳 30 ～60 min。

步骤 2：准备 DNase I 足迹分析缓冲液

缓冲液的用量须根据研究的规模，依经验而定。

1. 解冻必要数量的蛋白质稀释缓冲液等份。

2. 准备 DNase I 稀释缓冲液：向 500 μl 蛋白质稀释缓冲液中加入 20 μl 1 mol/L $CaCl_2$。置于冰上或 4℃放置。

   加入 $CaCl_2$ 是为了增强 DNase I 的活性。当对提取物或不纯的蛋白质制剂进行足迹分析时，加入 $CaCl_2$ 非常必要。

3. 将储液稀释为终止缓冲液。

4. 使用前（见步骤 4：2），立即将实验所需用量的终止缓冲液分成小等份，并加入 1/1000 体积的 10 mg/ml 的蛋白酶 K 溶液。

步骤 3：DNA 结合反应

该步骤用于滴定起始 DNase I 和 DNA 结合蛋白的浓度。较高浓度的 DNA 结合蛋白，特别是粗提物，倾向于抑制 DNase I 的活性。因此，用于裸 DNA 对照的 DNase I 浓度应小于用于高浓度 DNA 结合蛋白的 DNase I 的浓度。此外，随着片段大小的增加，所需 DNase I 的量减小。

1. 在冰上准备反应混合液：

DNA 模板（50 fmol/μl）	0.10 μl
poly（dI-dC）（1 μg/μl）	0.20 μl
BSA（50 μg/μl）	0.20 μl
β-巯基乙醇（14 mol/L）	0.10 μl
$MgCl_2$（0.1 mol/L）	1.5 μl
缓冲液 D	11.5 μl
H_2O	5.4 μl
总体积	19.00 μl
一批 14 个反应	266.0 μl

   配制多出两个反应量的足够的混合液，一份用于不可避免的损失或移液过程中的不一致性，另一份用于无 DNase I 的模拟反应。对于粗提物，从上述混合液中除去缓冲液 D 和（或）H_2O，这样可以添加 12～17 μl 的提取物（见步骤 3：5）。同时要注意在该体系中会含有内源性核酸酶的粗提物。在这些情况下不加 $MgCl_2$ 的结合反应可能效果更好，$MgCl_2$ 可以加到后面的 DNase I 稀释缓冲液中。

2. 在冰上向 13 个 0.5 ml 的硅化 Eppendorf 管，每管分装 19 μl 反应混合液。

3. 立即滴定前在冰上融化重组 DNA 结合蛋白。

4. 取 1 μl 融化的重组 DNA 结合蛋白加入 39 μl 蛋白稀释缓冲液中，配制一份“低浓度”样本。手指轻弹管壁温和混匀。

不推荐使用自动移液器剧烈混匀，因为有些蛋白质可能对该操作中产生的泡沫敏感。

5. 向反应管中加入 DNA 结合蛋白。

管 1～4：1 μl 未稀释的 DNA 结合蛋白

管 5～8：1 μl “低浓度” DNA 结合蛋白

管 9～12：1 μl 蛋白稀释缓冲液

管 13（“无 DNase I”对照）：1 μl 未稀释的 DNA 结合蛋白

对于粗提物，必须以低量和高量的提取物滴定 DNase I。注意粗提物是 DNase I 的强抑制剂，必须加入大量 DNase。另外，对于粗提物（或如果蛋白质浓度特别稀），可加入多达 17 μl 的提取物（见步骤 3：1）。

6. 将各管置于 30℃的水浴中孵育 20 min（对于粗提物，置于室温或冰上孵育）。

虽然有些蛋白质（如 TBP）的结合较为缓慢，但这个反应时间通常足以使 DNA 结合蛋白的结合达到平衡。

步骤 4：DNase I 足迹

此处的目的是对于给定量的 DNA 结合蛋白，确定产生均匀分布的切割梯带，并且确定 50%的 DNA 仍未被切割时的 DNase I 的浓度。这会将单个 DNA 分子内的多次切割降到最低；单个 DNA 分子内的多次切割会影响数据的解释。

1. 用 DNase I 稀释缓冲液（来自步骤 2：2）将 DNase I 系列稀释到 1/9、1/27、1/81和 1/243。

所需的浓度高度取决于 DNase I 的来源、混合物中二价阳离子的浓度及所用的 DNA 结合蛋白。注意对于比此处所用的片段更大的片段，将需要较小的 DNase I 用量。

2. DNA 结合孵育 20 min 后（步骤 3：6），按如下次序进行 1 min DNase I 消化。

时间（min：sec）	管号	DNase I（稀释）	终止缓冲液
0：00	1	1 μl（1/9）	—
0：15	2	1 μl（1/27）	—
0：30	3	1 μl（1/81）	—
0：45	4	1 μl（1/243）	—
1：00	1	-	100 μl
1：15	2	-	100 μl
1：30	3	-	100 μl
1：45	4	-	100 μl

加入 DNase I 后要立即轻弹管壁使之温和混匀。用粗提物时，与结合反应一样，DNase I 消化反应在冰上进行可能效果更好。

最好将 4 份 DNase I 消化缓冲液紧放在要进行滴定的管后。旁边放一个放射性废弃物处理容器，因为移液器的末端会有轻微的放射性。

3. 尽快对管 5～8 和管 9～12 重复步骤 4：2。

4. 向管 13 中加入 1 μl DNase I 稀释缓冲液（不含 DNase）进行模拟反应，1 min 后再加入 100 μl 终止缓冲液。

在此模拟反应中产生的未被切割的探针可与上面产生的切割梯带作比较。

5. 将所有的反应管于 55℃孵育 15 min，使蛋白酶 K 消化发生。

对于较粗的蛋白质制剂，这一步骤尤为重要。

6. 孵育样品的过程中，将聚合好的凝胶装到凝胶电泳装置上，1000 V 预电泳 30 min。

7. 用等体积（120 μl）酚抽提每份反应混合液。

8. 将所有管离心 2 min。

9. 转移恒定体积（如 100 μl）的上层水相至 1.5 ml 未硅化的 Eppendorf 管中。

10. 等体积（100 μl）酚∶氯仿重复步骤 4：7～4：9。

11. 往 1.5 ml 管中加入 2 倍体积的冰冷的 95％乙醇以沉淀 DNA，充分混匀。

没有必要加入 3 mol/L 乙酸钠，因为终止缓冲液含有高盐。

12. 所有管置于干冰、干冰/乙醇或－80℃冰箱冷冻 10 min。

13. 在微量离心机中将所有管 14 000 *g* 离心 10 min。

14. 利用细长的窄口巴斯德吸管从沉淀中吸除乙醇。

沉淀疏松，要小心！用盖革计数器定期监测试管，确保沉淀未偶然吸入巴斯德吸管。

15. 用 100 μl 80％乙醇洗涤沉淀，离心 2 min，用细长吸管除去乙醇，使沉淀在空气中彻底干燥。

步骤 5：凝胶电泳和放射自显影

1. 用甲酰胺染色混合液重悬沉淀。
2. 95℃孵育样品 2 min。
3. 点样至聚丙烯酰胺/尿素凝胶上，进行凝胶电泳。
4. 真空干燥凝胶。
5. 进行放射自显影或磷屏成像分析。

步骤 6：数据分析

DNase I 滴定的目的是优化切割条件，以便得到均匀分布的切割梯带，同时至少有 50％的探针维持未被切割的状态。为了精确地估计切割量，必须将消化梯带和仅含探针、不含 DNase I（如管 13）的模拟反应物进行比较。用粗蛋白提取物进行反应时，模拟反应还可显示内源性核酸酶的存在。理想情况下，随着 DNase I 浓度的增加，完整探针的量应该随之减小，而切割产物的量应该增加。当探针开始消失时，切割产物的分布应该会发生改变，从低迁移率的条带逐步变成高迁移率的条带，因为 DNase I 开始将探针切割成更小的片段。理想情况下，应该记录未稀释的和稀释 40 倍的 DNA 结合蛋白都生成大约 50％切割时的 DNase I 的用量，以用于计算剂量应答曲线中的用量（见下文步骤 7）。

步骤7：绘制剂量应答曲线

虽然没有必要为每个蛋白质浓度校正DNase I用量，但在蛋白质处于可能产生抑制作用的浓度时，应基于未稀释样本和稀释样本之间抑制作用的差异，校正DNase I的用量。假定抑制作用与DNase I浓度呈某种线性关系，并由此推算用量。

1. 用预制的丙烯酰胺凝胶混合液倒胶，并按照上文所述（步骤1和步骤2）准备DNase I足迹分析缓冲液。

2. 用蛋白质稀释缓冲液配制DNA结合蛋白的8个3倍系列稀释：1（如未稀释的）、1/3、1/9、1/27、1/81、1/243、1/729和1/2187。

   对于粗提物，应滴定1～17 μl。

3. 在冰上制备一批足够用于10个反应的混合液（终体积190 μl；见步骤3：1）。

4. 在冰上向9个硅化0.5 ml Eppendorf管每管分装19 μl反应混合液。

5. 每管加入1 μl蛋白质稀释物。准备1管不加蛋白质（管9）的反应，用作对照。

6. 将所有管置于30℃水浴锅中孵育20 min（用粗提物时，置于室温或冰上孵育）。

7. 向每个反应中加入最佳用量的DNase I（根据步骤6中确定的量）。按照上述操作进行消化并处理产物（步骤4：2～5：5）。

步骤8：剂量应答数据分析

理想情况下，足迹应随着DNA结合蛋白浓度的增加而变得逐渐明显。占据比例或保护百分比是亲和计算中非常有用的两个术语，可以通过使用激光光密度法或磷屏成像软件将“无蛋白”泳道中切割梯带的条带和足迹泳道中的条带相比较，以此来推断占据比例或保护百分比。足迹上下的未影响条带可用于归一化DNase I切割效率。大多数蛋白质会与DNA位点逐渐结合，符合正常的一级反应米氏动力学。但是，许多蛋白质以二聚体或更高的寡聚体的形式结合，如果二聚化常数接近该位点的K_d，蛋白质结合会符合二级或更高级的反应动力学。如果该项技术要定量地用于建立DNA结合动力学模型，关于技术考虑的更详细的概述见Koblan等（1992）。

附加说明

该方案的通用化

没有用于通用化该方案的一成不变的规则，使用时应该考虑以下因素。

· 较小的反应规格将有利于不太浓缩的DNA结合蛋白。

· 较低的盐浓度将有利于结合较弱的蛋白质。

· 终体积可以是10～100 μl之间的任何体积，这将取决于反应。较小的反应体积可用于节约有价值的材料。

· 不同的载体DNA会有不同的效应：如poly（dI-dC）可用于竞争TBP和TATA盒探针的结合，常用poly（dG-dC）代替。

关于相关参数的详细分析，以及这些参数如何影响 DNA-蛋白质相互作用，见综述 Jen-Jacobsen（1997）、Record 等（1991）及 Koblan 等（1992）。下面是发表的著作中已被成功使用的一些参数的变动范围。

· 10～50 mmol/L HEPES-KOH 或 Tris-HCl（pH7.0～8.0）。DNase I 兼容范围很宽的缓冲液。

· 0～10 mmol/L $MgCl_2$。这种二价阳离子中和磷酸，并且已报道有些 DNA 结合蛋白与其结合作为一种辅助因子。

· 50～100 mmol/L KCl、NaCl、乙酸钾、硫酸铵。虽然大多数 DNA 结合蛋白对异常高浓度的盐都敏感，但最佳盐浓度会依赖于蛋白质显著变化，最好根据经验进行评估。

· 0～20%（*V/V*）甘油。甘油是稳定剂，可以降低反应中水的浓度，模拟体内环境，它也是一种自由基清除剂。

· 10～100 μg/ml BSA。BSA 是另一种蛋白质稳定剂，可作为非特异性载体，可以阻碍稀释蛋白样品对表面（如 Eppendorf 管壁）的假结合。

· 0.01%～0.1% NP-40 或 TritonX-100。这些非离子型去污剂阻止蛋白质与表面的非特异性结合，也可作为抗聚合剂。

· 0.1～1 mmol/L DTT 或 10～50 mmol/L β-巯基乙醇。这些还原剂对许多蛋白质来说是绝对必不可少的。

· 0～1μg 载体 DNA、小牛胸腺 DNA 或合成的共聚物（dI-dC）。阻止结合反应中杂质的非特异性结合，最小化足迹分析蛋白质的非特异性结合。

· 聚乙烯醇（PVA）、聚乙二醇（PEG）、二甲基亚砜（DMSO）。PVA 和 PEG 试剂的浓度可以显著变化。PVA 和 PEG 是体积排阻剂，能够增加蛋白质在溶液中的有效浓度，减小水的浓度。DMSO 是一种变性剂，在低浓度时具有一些不寻常的稳定特征，可能是最小化蛋白质的非特异性结合或有利于蛋白质构象的柔韧性。

DNA 片段大小

关于片段大小，有几个重要考虑。当使用 DNA 结合机制或亲和力已经确定的高度纯化的蛋白质时，应使用小的 DNA 片段，如克隆到载体 pGEM 或 pUC 多聚接头中的位点（长度为 50～100 bp）。小的 DNA 片段在聚丙烯酰胺/尿素测序胶上能产生更好的条带分辨率，而且可以将同一片段在 EMSA 和足迹研究中相比较。当要确定蛋白质结合位点在启动子上的位置时，必须使用大片段（见第 7 章）。在粗提物中，含有大量非特异性 DNA 末端结合蛋白，如 Ku 自身抗原。因此，目的位点最好距离末端 50 bp 左右，以便区别足迹和杂质性的末端结合物。注意对于比此处使用的片段更大的片段，将需要更小量的 DNase。

提取物

足迹研究中使用提取物增加了实验设计的复杂性。为克服大量的抑制物，必须使用较高浓度的 DNase I。有些提取物富含非特异性 DNA 结合蛋白。因此，为了最小化抑

制作用，必须使用低浓度的提取物（小于 10 μg）和高浓度的非特异性载体 DNA（1 μg 或更多），这必须与检测特异性结合时对高浓度提取物的需要相平衡。由于内源性核酸酶的存在，有时谨慎的做法是在低温（4℃）条件下进行最初的结合及 DNase Ⅰ反应。但是，较低温下的切割需要较高浓度的 DNase Ⅰ，用量根据经验而定。或者为抑制结合反应过程中的核酸酶活性，预孵育时不加 $MgCl_2$，之后，$MgCl_2$ 随 DNase Ⅰ加入。

时间进程

有时有必要测量蛋白质结合的动力学。理想情况下，这意味着蛋白质加入结合混合液的时间和混合液置于适当的温度下孵育的时间要错开。简单地在冰上将蛋白质加入混合液中并假定在 30℃孵育之前它不会与 DNA 结合是不正确的。如上文所提到的，许多结合研究都是在冰上进行的。因此，理想情况下，应将反应混合液置于孵育温度下，然后加入 DNA 结合蛋白。

事实上，有些反应达到平衡的速度非常快（如 GAL4-VP16 的结合在 2 min 内完成），而其他反应则缓慢（TBP 的结合需要 1 h 或更长时间）。在前者情况下，谨慎的做法是通过增加 DNase Ⅰ浓度和将切割时间降低到约 15 s，从而大大加快 DNase Ⅰ切割反应。相反，对于慢反应，可以通过降低 DNase Ⅰ浓度放缓其切割的速度。其他研究者已评论了建立生物物理模型时进行真正的定量足迹分析所需的条件（Brenowitz et al. 1986）。

使用非标记 DNA

可以对环状 DNA 模板进行 DNase Ⅰ足迹分析，并通过 LM-PCR 或间接末端标记来检测切割点（Gralla 1985；Grange et al. 1997）。

疑难解答

无切割

可能原因：核提取物对于商业化来源的 DNase Ⅰ浓度有强烈抑制作用。

解决办法：如果没有观察到 DNase Ⅰ的切割，可以增加 DNase Ⅰ、$MgCl_2$ 或 $CaCl_2$ 的浓度，或试用不同批号的 DNase Ⅰ。尝试降低提取物浓度或通过柱层析对其进一步分级分离。另外，还可以购买 DNase Ⅰ粉末，配制成高浓度的 DNase Ⅰ。

无或弱足迹

可能原因：这可能归因于蛋白质浓度低、片段中缺乏生理性位点或蛋白质制剂中掺杂结合抑制剂。

解决办法：提高蛋白质浓度、降低 DNA 浓度或调整结合参数，使盐浓度、pH 等达到最佳。

拖尾凝胶

可能原因：拖尾凝胶经常是由以下原因造成的。①未预电泳凝胶除去丙烯酰胺中残

留的过硫酸铵；②酚抽提步骤中蛋白质去除不彻底，之后与乙醇共沉淀，使产生的沉淀难以重悬；③沉淀物中含有太多的盐。

解决办法：在情况①中，条带在盐前处被压缩，下一次实验前，对凝胶进行预电泳；在情况②中，可以观察到从点样孔向下在凝胶上弥散的黑点，尝试更加剧烈地悬浮样品，或乙醇沉淀前对样品进行再次酚抽提，另一种做法是利用 2 mol/L 乙酸铵沉淀样本，乙酸铵沉淀倾向于使蛋白质、小分子寡核苷酸和核苷酸保留在溶液中，而使较大的 DNA 片段得以沉淀；在情况③中，通常观察到反向楔形效应，即靠近凝胶底部的条带变窄，干燥前用 80%的乙醇简单洗涤沉淀，然后重悬于甲酰胺染色缓冲液中。

方案 11.2　羟自由基足迹

羟自由基足迹法是由 Tullins 和 Dombroski 于 1986 年创立的，最初用于研究λ抑制因子和 Cro 与 DNA 的结合。现阶段，该方法已广泛应用于研究蛋白质-DNA 相互作用和蛋白质结合过程中发生在 DNA 上的结构扰动（如弯曲）。羟自由基通过从小沟脱氧核糖的不同氢原子上抽取氢（$5'H > 4'H > 3'H = 2'H = 1'H$），切割 DNA（Balasubramanian et al. 1998）。蛋白质在小沟部位的结合一般保护糖分子免受切割（见专题 11.4 和图 11.11），该方法可提供蛋白质与小沟结合的详细信息。虽然有些信息与从乙基化干扰分析得到的信息相重叠，但它通常具有互补性，有助于建立蛋白质如何与其位点对接的模型。自由基由 Fe（II）EDTA 产生，Fe（II）EDTA 将过氧化氢切割成羟自由基和氢氧根离子，自由基随即在有限度的扩散反应中切割糖分子。加入抗坏血酸是为了再生活性 Fe（II）。

羟自由基具有高度反应性，但易被甘油灭活，因此，虽然甘油是大多数结合反应的组分之一，但在这种情况下必须避免加入甘油。羟自由基反应需要过氧化氢，它是一种强氧化剂，偶尔会干扰蛋白质与 DNA 的结合，可以偶尔调整反应条件来适应这样的情况。自由基也和蛋白质反应，并切割蛋白质，反应时间过长会导致蛋白质降解。

时线与安排

除了缓冲液中不加甘油，该反应体系的建立几乎与 DNase I 足迹相同。通常，DNA 质粒的品质对该方法的成功很关键。许多制剂中都事先已存在低水平 DNA 切割。因为羟自由基的衰变一般相当迅速，在该过程中通常只有不足 5%的 DNA 被切割。因此，少量事先已存在的背景切割会引起信噪比问题。使用经 CsCl 密度梯度离心纯化两次的高质量质粒 DNA 很重要。而且，当使用碱裂解法时，异丙醇沉淀后的 DNA 团块在 CsCl 密度梯度离心之前必须用 Tris 碱中和。一旦得到了高质量 DNA，就可以制备末端标记的 DNA 片段。通常有必要在羟自由基反应条件下进行一次 DNase I 足迹分析，以确定目的蛋白质在这些条件下与位点相结合。实际上，如果蛋白质中含有甘油之类的抑制剂，在蛋白质的不同浓度下滴定自由基切割时间可能是有用的。一旦这些准备步骤完成，就可以在一天内完成羟自由基足迹分析，次日就可通过放射自显影或磷屏成像仪分析产物。

提纲

步骤1：配制缓冲液（30 min）。
步骤2：羟自由基保护（5～6 h）。
步骤3：凝胶电泳和放射自显影。
步骤4：分析数据。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

抗坏血酸（10 mmol/L）
结合缓冲液
  20 mmol/L HEPES-KOH（pH 7.5）〈!〉
  50 mmol/L KCl〈!〉
  5 mmol/L $MgCl_2$〈!〉
  100 μg/ml 牛血清白蛋白（BSA）
  5 μg/ml poly（dI-dC）
^{32}P标记的DNA〈!〉
DNA结合蛋白
干冰〈!〉
EDTA（200 μmol/L）
乙醇（80%和95%）〈!〉
Fe $(NH_4)_2SO_4$（100 μmol/L）〈!〉
甲酰胺染色混合液
  98% 去离子甲酰胺〈!〉
  10 mmol/L EDTA（pH 8）
  0.025% 二甲苯青FF〈!〉
  0.025% 溴酚蓝〈!〉
H_2O_2（0.3%）〈!〉
苯酚〈!〉（可选）
聚丙烯酰胺/尿素凝胶混合液〈!〉
终止液
  5 mol/L 乙酸铵
  5 mmol/L 硫脲〈!〉
  10 mmol/L EDTA
  100 μg/ml tRNA

设备

电泳仪
Eppendorf 管（1.5ml）
−80℃冰箱（可选）
凝胶干燥机，真空
微量离心机
巴斯德吸管（窄口）
磷屏成像仪暗盒
水浴
X 光片（可选）

程序

步骤 1：配制缓冲液

1. 为每个反应新鲜配制结合缓冲液、终止液和甲酰胺染色混合液。

   对于终止液，从浓缩储液中新鲜稀释硫脲。

2. 自贮液稀释 Fe（NH_4）$_2SO_4$、EDTA、抗坏血酸及 H_2O_2 溶液到上面标明的浓度。

   不要在玻璃器皿中稀释，使用 Eppendorf 离心管。

步骤 2：羟自由基保护

对于每组实验，包括一组 3 个对照样本并平行孵育，每个样本中分别去除 Fe-EDTA（步骤 2：3）、抗坏血酸（步骤 2：4）或过氧化氢（步骤 2：5）。

1. 往 50～100 μl 结合缓冲液中加入 10～50 fmol ^{32}P 标记的 DNA 片段和饱和浓度的重组 DNA 结合蛋白。
2. 室温下（或在 DNase I 足迹分析确定的最适温度下）孵育 15 min。
3. 加入 10 μl 新鲜稀释的 100 μmol/L Fe（NH_4）$_2SO_4$ 和 10 μl 新鲜稀释的 200 μmol/L EDTA。
4. 加入 10μl 10 mmol/L 抗坏血酸。
5. 加入 10 μl 0.3% H_2O_2 溶液，室温下孵育 2 min。
6. 加入 100 μl 终止液终止反应。
7. 在继续后续实验之前酚抽提（可选）。

步骤 3：凝胶电泳和放射自显影

1. 乙醇沉淀 DNA。用 80%乙醇洗涤沉淀，干燥后重悬于甲酰胺染色混合液中。
2. 95℃孵育 2 min，变性 DNA。
3. 直接上样到 10%或 12%聚丙烯酰胺/尿素凝胶上。根据片段大小电泳 4～6 h。

4. 真空干燥凝胶。
5. 将干燥的凝胶在 X 光片或磷屏成像仪上曝光。

附加说明

羟自由基应该产生在凝胶上曝光强度大致相等的 DNA 片段梯带（除了发生弯曲或扭转的 DNA 区段）。梯带的产生应依赖于过氧化物、抗坏血酸和 Fe-EDTA 的加入。如前文所述，有些蛋白质在过氧化氢存在时不能结合。可以轻松操控过氧化氢、抗坏血酸和 Fe-EDTA 的相对平衡，以适应替代反应条件（如降低过氧化氢的浓度、增加抗坏血酸的量）(Dixon et al. 1991)。为了确定蛋白质是否受到自由基条件的影响，可以在产生自由基的化学试剂存在的条件下，进行 DNase I 足迹反应。如果在足迹分析中某些化合试剂及其组合抑制结合，则必须相应地调整反应条件。自由基也可以检测 DNA 的改变（如弯曲），这是因为这些改变的特点是一系列交替排列的切割减弱和切割增强的 5 bp 区域，这种现象是由 DNA 弯曲上小沟的变窄或扩大造成的。自由基还可检测其他异常的 DNA 结构（Dixon et al. 1991）。

这项技术有多种变种，包括核苷缺失实验。首先用羟自由基处理 DNA，使得平均每个分子上随机去除一个核苷。依靠 EMSA 测定蛋白质结合有缺口 DNA 的能力。对结合 DNA 和未结合 DNA 进行剪切、纯化，并在聚丙烯酰胺/尿素凝胶上分级分离。结合级分富含缺失了对蛋白质结合不重要的碱基的分子，而未结合级分富含缺失了对蛋白质结合重要的碱基的分子。通过与该片段的测序梯带在聚丙烯酰胺/尿素凝胶上一同电泳，可鉴定这些碱基的位置。该项技术可用于在单个实验中以单核苷酸分辨率扫描 DNA 分子的研究（Hayes and Tullius 1989）。

疑难解答

无切割

可能原因：抗坏血酸、过氧化物和 EDTA 稀释液都有可能在短时间内失活，即使是干燥的试剂，特别是抗坏血酸，也会在潮湿的气候中变质。

解决办法：如果未观察到切割梯带，尝试使用新鲜化学试剂。另外，要保证不管是 DNA 结合蛋白制剂还是缓冲液中都不含有多于 1%的甘油。然而，即使 1%的甘油也具有抑制性。

无特异性切割

可能原因：DNA 制备不够仔细。

解决办法：如果在缺乏切割的情况下就事先已存在梯带，制备质粒 DNA 或 DNA 片段时就应倍加小心，尝试重新制备 DNA，还要配制新鲜的去离子甲酰胺上样染色混合液。

方案 11.3 磷酸乙基化干扰分析

乙基化干扰分析最初是由 Walter Gilbert 及其同事发展的，用于研究 RNA 聚合酶与 *E. coli* 启动子的结合（Siebenlist and Gilbert 1980；Siebenlist et al. 1980）。该技术（综述，见 Manfleld and Stockley 1994）利用乙基亚硝基脲，将含有一个蛋白质识别位点的^{32}P末端标记 DNA 中的磷酸乙基化。这一修饰中和了磷酸的电荷，并沿小沟边缘放置了一个庞大的乙基基团（见图 11.12）。大多数位点旁侧的磷酸不干扰（见图 11.14），只有位于位点中的一小部分磷酸会干扰，但是会产生干扰的那些磷酸干扰蛋白质靠近结合位点的小沟或磷酸骨架。

然后通过 EMSA（或滤膜结合、免疫沉淀等）分离结合 DNA 和未结合 DNA。对凝胶进行放射自显影，并从凝胶上切下对应于结合和游离 DNA 的条带，并从凝胶切片中纯化 DNA。未结合级分富含修饰作用发生在干扰蛋白质结合的位置上的 DNA 分子。相比之下，结合级分的修饰作用发生在不干扰蛋白质结合的位置上。乙基化还使得 DNA 对碱诱导的切割敏感。之后，将切割产物和测序梯带一同在聚丙烯酰胺/尿素凝胶上分级分离。注意以强碱切割会使磷酸留在 5′ 或 3′ 断点上。因此，切割产物通常在凝胶上以紧密间隔的双带迁移。

时线与安排

乙基化干扰分析可以在两天内完成。大多数安排首先涉及放射性标记片段，并通过 EMSA 确定几乎饱和的蛋白质用量（方案 11.5 和方案 11.6）。在最初的 EMSA 中使用较大量的 DNA，以确保实际程序中 DNA 的充分回收所需的条件。将 DNA 片段乙基化、纯化并用于另一轮的 EMSA；EMSA 后将结合 DNA 和未结合 DNA 从胶上切出并纯化。纯化的 DNA 用哌啶切割，并在聚丙烯酰胺/尿素测序胶上与测序梯带一同分级分离。

提纲

步骤 1：配制缓冲液（1 h）。
步骤 2：进行乙基化干扰分析（10 h）。
步骤 3：切割乙基化的残基（1 h）。
步骤 4：在聚丙烯酰胺/尿素凝胶上分级分离 DNA，分开结合级分和未结合级分（3 h）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

乙酸铵（5 mol/L）
β-巯基乙醇（14 mol/L）〈!〉

牛血清白蛋白（BSA；50 μg/μl）

缓冲液 A

　10 mmol/L Tris-HCl（pH 8.0）

　10 mmol/L $MgCl_2$〈!〉

　100 mmol/L KCl〈!〉

　1 mmol/L DTT（二硫苏糖醇）〈!〉

　0.1 mmol/L EDTA

缓冲液 D

　20 mmol/L HEPES-KOH（pH 7.9）〈!〉

　0.2 mmol/L EDTA

　20% 甘油

　0.1 mol/L KCl

切割缓冲液

　10 mmol/L Na_3PO_4（pH 7.0）

　1 mmol/L EDTA

^{32}P 标记的 DNA〈!〉

DNA 结合蛋白

乙醇（80%和 95%）〈!〉

乙基亚硝基脲〈!〉

甲酰胺染色混合液

　98% 去离子甲酰胺〈!〉

　10 mmol/L EDTA（pH 8）

　0.025% 溴酚蓝〈!〉

　0.025% 二甲苯青 FF〈!〉

$MgCl_2$（0.1 mol/L）

NaOH（1 mol/L）〈!〉

聚丙烯酰胺凝胶混合液，非变性〈!〉

关于聚丙烯酰胺混合及灌注的说明，见 Sambrook 和 Russell（2001）。

聚丙烯酰胺/尿素凝胶混合液〈!〉

Poly（dI-dC）（1 μg/μl）

二甲砷酸钠（0.05 mol/L，pH 8.0）〈!〉

设备

电泳仪

Eppendorf 管（1.5 ml）

－80℃冰箱（可选）

真空凝胶干燥机

微量离心机

巴斯德吸管（窄口）
磷屏成像仪暗盒
水浴
X 光片（可选）

程序

步骤 1：配制缓冲液

1. 为每个实验新鲜配制缓冲液 A、切割缓冲液及二甲砷酸钠溶液。
2. 灌注并预电泳 4.5％的聚丙烯酰胺凝胶。

步骤 2：乙基化干扰分析

1. 向 10 pmol ^{32}P 末端标记 DNA 片段中加入 100 μl 二甲砷酸钠（ pH 8.0）和 100 μl用乙基亚硝基脲饱和的 95％乙醇。
2. 50℃孵育 1 h。
3. 用 10 μl 5 mol/L 乙酸钠和 2.5 倍体积的乙醇沉淀 DNA 两次。
4. 用 100 μl 缓冲液 A 重悬沉淀。
5. 在冰上，在标准结合条件下（见方案 11.1)，将饱和浓度的 DNA 结合蛋白加入 50 000 cpm 的乙基化 DNA 片段中，30℃孵育 15 min。
6. 小心地将样品载样到已预电泳的凝胶上，进行电泳。
7. 使凝胶在胶片上曝光。
8. 将凝胶和放射自显影胶片对齐，用剃刀刀片切下与结合级分和未结合级分对应的条带。

   确保凝胶和胶片的方向正确（从聚丙烯酰胺凝胶上分离放射性片段的方法，见第 3 章，方案 3.2，步骤 3)。
9. 通过电洗脱从凝胶上回收 DNA 片段（方案 11.6，步骤 3)。
10. 乙醇沉淀，并干燥 DNA 沉淀。

步骤 3：切割乙基化的残基

1. 用切割缓冲液重悬来自结合级分和未结合级分的沉淀。
2. 加入 2.5 μl 的 1 mol/L NaOH，90℃孵育 30 min 。

   在该步中，乙基化的 DNA 骨架被 NaOH 切除。
3. 样品按 1∶1 比例加入甲酰胺染色混合液。将样品和测序梯带一同加样到 10％的聚丙烯酰胺/尿素凝胶上，跑胶。
4. 真空干燥凝胶。
5. 使凝胶在 X 光片上曝光，进行放射自显影或磷屏成像分析。

附加说明

必须根据经验确定最初的乙基化频率（图 11.13)。改变乙基亚硝基脲的用量和

（或）修饰时间。分析裸 DNA 分子的切割，确定每个 DNA 上少于一个修饰磷酸的条件（专题 11.4）。

在干扰技术中，对于提取有用的信息而言，100%的结合是不必要的。结合泳道虽然富含不干扰结合的乙基化磷酸，但乙基化磷酸被耗尽，这些乙基化磷酸确实会干扰结合。将来自结合泳道的被切割 DNA 与标准乙基化梯带直接比较，并确定结合泳道中缺失条带的位置，可以取得重要的信息。乙基化条件和直接结合实验不相容，因此没有乙基化保护程序。从乙基化干扰获得的数据通常可有助于确定骨架接触，并且如方案 11.2 中所讨论的，可以和羟自由基数据作比较。

疑难解答

加入蛋白质时乙基化的 DNA 不发生凝胶迁移

可能原因：这可能指示 DNA 高度过乙基化。

解决办法：尝试更低浓度的乙基亚硝基脲或更短的修饰时间。

可能原因：DNA 中可能没有充分从乙基亚硝基脲中纯化。

解决办法：用乙醇重新沉淀 DNA 一次或两次，保证重悬 DNA 的 pH 接近 7。

当在聚丙烯酰胺/尿素凝胶上分离并比较时，来自 EMSA 的结合 DNA 级分和未结合 DNA 级分表现出同样的条带分布。

可能原因：这一结果通常指示来自不纯或陈旧 DNA 的事先存在的条带或切割模式。

解决办法：制备新鲜的质粒制剂，参照羟自由基足迹分析程序中所列的注意事项（方案 11.2）。

方案 11.4　甲基化干扰分析

甲基化干扰和保护技术是由 Walter Gilber 及其同事发展的，用于研究 *E. coli* RNA 聚合酶和 DNA 大沟相互作用。该方法在化学上是 Maxam 和 Gilbert 发明的化学法 DNA 测序所使用的化学技术的一个分支。甲基化现在被广泛地作为一种研究蛋白质与大沟中鸟嘌呤相互作用的主要技术，尽管它也能检测小沟中的腺嘌呤和单链 DNA 中的胞嘧啶。

该方法有两种形式。在保护技术中，蛋白质预先与含有识别位点的^{32}P 末端标记 DNA 结合，然后用 DMS 攻击蛋白质-DNA 复合物。结合的蛋白质保护大沟中的鸟嘌呤和小沟中的腺嘌呤免受甲基化和随后的哌啶切割。虽然保护技术对于许多蛋白质都平稳起效，但是它具有几个缺点。因为 DMS 修饰某些缓冲液（如 Tris），历史上保护反应都是在二甲砷酸盐中进行的，该试剂如今已很少用作标准蛋白质结合缓冲液。此外，DMS 直接修饰有些蛋白质，这些修饰可扰乱蛋白质与 DNA 的结合（如 TFIIIA）。而且，和干扰技术不同，为了观察到足迹并获得结合位点的相关信息，保护技术几乎需要化学计量的结合（约 90%）。

由于这些原因以及其他下文所述的原因，甲基化干扰技术（改编自 Ausubel et al. 1994，p. 12. 3. 1）已被广泛采用，并取代了甲基化保护方法，它无疑是用于检测参与蛋白质序列特异性识别碱基的分辨率最高的方法之一。该项技术涉及含有识别位点的^{32}P 末端标记 DNA 的修饰（平均每个 DNA 分子修饰一次，见图 11. 14），以及之后的蛋白质结合。某些修饰阻止蛋白质结合，因为这些修饰位于蛋白质与 DNA 紧密靠近的位置，但是大多数修饰对结合没有影响。然后，通过 EMSA（或滤膜结合、免疫沉淀等）分离结合 DNA 和未结合 DNA。对凝胶进行放射性自显影，并从凝胶上切下对应于结合和游离 DNA 的条带，再从凝胶切片中纯化出 DNA。未结合级分将富含其修饰作用发生在干扰蛋白质结合的位置上的 DNA 分子。相比之下，结合级分将含有不干扰蛋白质结合的修饰。甲基化削弱核苷酸，并使其对哌啶诱导的脱嘌呤作用敏感，这反过来引起 DNA 骨架的切断。之后，被切割的片段在测序胶上与 A+G 和 C+T 化学测序梯带一同分级分离，以鉴定受影响的位置。

时线与安排

DMS 干扰分析可以在 2 天内完成。第 1 天，DNA 探针进行甲基化，和目的蛋白一起进行 EMSA，分离并纯化。第 2 天，利用哌啶切割甲基化的 DNA 探针，在测序胶上对切割片段进行分析，然后进行放射自显影或磷屏成像分析。

提纲

步骤 1：配制缓冲液（1 h）。

步骤 2：制备甲基化的 DNA 探针（2 h）。

步骤 3：使甲基化的探针与蛋白质结合，并通过 EMSA 分离游离 DNA 和结合 DNA（5 h）。

步骤 4：用哌啶切割 DNA（3 h）。

步骤 5：在测序胶上分析切割梯带（4 h）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

乙酸盐/EDTA

3 mol/L 乙酸钠	5 ml
0. 5mol/L EDTA	10 μl
dH_2O	到 50 ml

将 pH 调到 5. 2。

β-巯基乙醇（14 mol/L）〈!〉

牛血清白蛋白（BSA；50 μg/μl）

缓冲液 D

20 mmol/L HEPES-KOH（pH 7.9）〈!〉
20% 甘油
0.2 mmol/L EDTA
0.1 mol/L KCl〈!〉
硫酸二甲酯（DMS）〈!〉
DMS 反应缓冲液
50 mmol/L 二甲砷酸钠（pH 8）〈!〉
1 mmol/L EDTA（pH 8）
DMS 终止缓冲液
1.5 mol/L 乙酸钠（pH 7）
1.0 mol/L β-巯基乙醇
^{32}P 末端标记的 DNA 探针〈!〉
末端标记 DNA 的详细方案见方案 11.6。
DNA 结合蛋白
启动甲基化干扰分析开始前，必须先进行分析以优化饱和 EMSA 或 DNase 足迹反应（见方案 11.1 和方案 11.5）中未修饰 DNA 片段所需的蛋白质量。之后，DMS 干扰实验使用几乎饱和的蛋白质量。
干冰〈!〉
乙醇（80%和 95%）〈!〉
甲酰胺染色混合液
98% 去离子甲酰胺〈!〉
10 mmol/L EDTA（pH 8）
0.025% 溴酚蓝〈!〉
0.025% 二甲苯青 FF〈!〉
$MgCl_2$（0.1 mol/L）〈!〉
哌啶（1 mol/L）〈!〉
PMSF（0.1 mol/L）〈!〉
聚丙烯酰胺凝胶混合液（非变性）〈!〉
聚丙烯酰胺/尿素凝胶混合液〈!〉
Poly（dI-dC）（1 μg/μl）
TE 缓冲液
10 mmol/L Tris（pH 7.6）
1 mmol/L EDTA
tRNA（10 mg/ml）

设备

化学通风橱

电泳仪
Eppendorf 管（硅化，1.5-ml）
－80℃冰箱（可选）
微量离心机
针头（大口径）
巴斯德吸管
磷屏成像仪暗盒（可选）
闪烁计数器
真空蒸发器
振荡器
水浴
X 光片（可选）

程序

步骤 1：配制缓冲液

1. 用 dH_2O 配制 50 ml DMS 反应缓冲液，4℃保存。
2. 用 dH_2O 配制 50 ml 新鲜的 DMS 终止缓冲液和 50 ml 乙酸盐/EDTA。

步骤 2：制备甲基化的 DNA 探针

1. 将大约 10^6 cpm 的探针悬浮在 5～10 μl TE 缓冲液中。
2. 加入 200 μl DMS 反应缓冲液。
3. 在化学通风橱中加入 1 μl 未稀释的 DMS。涡旋振荡，使管内混合物充分混匀。
4. 室温孵育 5 min。
5. 反应中加入 40 μl DMS 终止缓冲液。
6. 加入 1 μl 10 mg/ml tRNA 和 600 μl 95%的乙醇，混匀。
7. 在干冰/乙醇浴中孵育 10 min。
8. 在微量离心机中 4℃下 14 000*g* 离心 10 min，用拉长的巴斯德吸管小心除去上清。

   将上清弃于液体 DMS 废物中。
9. 将沉淀物重悬于 250 μl 乙酸盐/EDTA 中。在冰上工作，加入 750 μl 95%乙醇，混匀，再次离心。
10. 重复前面的步骤（步骤 2：9），用拉长的巴斯德吸管小心除去上清。
11. 用 80%乙醇洗涤沉淀物，14 000 r/min 离心 10 min。用拉长的巴斯德吸管小心除去上清，使管完全风干。
12. 在闪烁计数器中对沉淀进行契伦科夫计数，确定 cpm。
13. 将沉淀以 20 000 契伦科夫 cpm/μl 的浓度重悬于 TE 缓冲液中。

    DNA 可能难以重悬。加热、涡旋及用吸管上下吹打，以使 DNA 溶解。

步骤3：甲基化探针结合蛋白质并分离结合和未结合探针

建立由3或4个DNA结合反应组成的反应系列，对反应个数略加扩大以增加cpm数。多个反应平行进行，以获得一致性的结果和增加的cpm数。

1. 对于每个反应，在硅化的0.5 ml Eppendorf管中混合下述组分：

DNA结合蛋白	1.00 μl
DNA探针（50 fmol/ μl）	0.50 μl
poly（dI-dC）（1 μg/μl）	0.20μl
BSA（50 μg/μl）	0.05μl
β-巯基乙醇（14 mol/L）	0.05μl
$MgCl_2$（0.1 mol/L）	0.75μl
PMSF（0.1 mol/L）	0.05μl
缓冲液D	6.6 μl
H_2O	3.8 μl
总体积	13.00 μl

2. 30℃孵育1 h。
3. 将结合反应物上样在4.5%非变性聚丙烯酰胺凝胶上电泳（见方案11.5）。
4. 对凝胶放射自显影。切出对应于蛋白质-DNA复合物（即结合）和游离探针的条带。

关于如何从聚丙烯酰胺上分离放射性片段的方法，见方案3.2，步骤3。

5. 通过电洗脱，从每个凝胶切片中纯化DNA（见方案11.6，步骤3）。

步骤4：用哌啶切割被修饰的DNA

1. 在化学通风橱中将来源于结合和未结合样品的沉淀物重悬于100 μl 1 mol/L哌啶中。
2. 将反应混合物于90℃孵育30 min。

用盖锁锁好管口，使试管不能爆开。

3. 孵育后，将管置于干冰上。
4. 用大头针在管的顶部扎孔，将样品置真空蒸发器中干燥1h，或直至干燥。
5. 加入100 μl蒸馏水，冷冻后再一次干燥，重复冷冻和干燥。检测样品的契伦科夫计数，以确定cpm。

步骤5：在DNA测序胶上分析片段

1. 根据契伦科夫计数，往沉淀物中加入足量的甲酰胺染色混合液，以使上样的1～2 μl样品中含3000 cpm。

2. 将样品置于 90℃热变性 5 min，然后在冰上骤冷。

3. 将游离探针和结合复合物样品上样到 6%～12%（根据片段大小而定）聚丙烯酰胺/尿素测序胶上。

　　对于利用增感屏的过夜曝光来说，每泳道 3000 cpm 一般足够，关键的是使结合和游离探针的计数相等，使样品之间能精确比较。

4. 在样品泳道旁上样一个 Maxam-Gilbert 或双脱氧测序梯带，以区分 G 和 A，并确定每种修饰的精确位置。

5. 在测序胶上电泳样品。使凝胶在 X 光片或磷屏成像屏上曝光过夜。

附加说明

记住在干扰技术中，对于获得有用的信息而言，100%的结合是不必要的。结合泳道尽管富含不干扰结合的甲基化 A 和甲基化 G，却耗尽会干扰的甲基化 A 和甲基化 G。直接比较来自结合泳道的被切割 DNA 和不用蛋白质而制备的标准甲基化梯带（不经 EMSA 分离），并确定缺失条带的位置（在结合泳道中），可取得重要的信息。

EMSA 之前 DNA 的严重过甲基化会导致单个 DNA 分子上多个位点发生甲基化，这会使该分析发生混乱，因为将不清楚究竟是哪个甲基化的碱基干扰蛋白质结合。通常基于带的曝光强度，可明显看出过甲基化，测序胶上缺乏未切割 DNA，或相对于切割产物只有少量未切割 DNA。

疑难解答

结合 DNA 和未结合 DNA 之间无差别

可能原因：从切割模式来看，甲基化效果好，但来自结合和游离样品的 DNA 产生相同的切割模式，有理由认为蛋白质结合区中没有 G。有些情况下，加入太多的蛋白质可克服甲基化 G 的负面影响，并且大多数 DNA 会出现在结合泳道中。

解决办法：降低滴定蛋白质的浓度，直到未结合 DNA 的量开始增加。

无鸟嘌呤切割

可能原因：DNA 可能从一开始就甲基化不足。

解决办法：滴定 DMS 用量和处理时间，优化修饰。

可能原因：哌啶可能已失效。

解决办法：用新鲜的新试剂再试一次。

事先存在切割梯带

可能原因：质粒 DNA 在制备期间可能已经广泛地脱嘌呤。

解决办法：制备新鲜的质粒 DNA，在 CsCl 梯度上纯化前，仔细确定 DNA 的 pH。

方案 11.5 电泳迁移率变动分析

在EMSA（或简称为“凝胶迁移”）中，将含有特定DNA位点的^{32}P标记DNA片段和相应的DNA结合蛋白一起孵育。通过非变性聚丙烯酰胺凝胶电泳将蛋白质-DNA复合物和游离（未结合）DNA分开。蛋白质阻滞与之结合的DNA片段的迁移。因此，游离DNA会比DNA-蛋白质复合物迁移得更快。用凝胶的放射自显影图像显示游离和结合放射性标记DNA的位置（Kerr 1995）。

EMSA是用于研究蛋白质对其位点结合特征的最灵敏的方法之一（见第7章）。该方法可用于推断一种蛋白质对一个或多个位点的结合参数和相对亲和力，也可用于比较不同蛋白质对同一位点的亲和力（Fried 1989）。还可用于研究含有多种蛋白质的更高级复合物。例如，一种蛋白质与一种位点的结合会产生一个占优势的迁移率变动复合物。如果另一种蛋白质结合在该蛋白质上，就会产生一个额外的迁移率变动复合物，或称“超迁移”复合物。最后，EMSA可用于研究蛋白质依赖性或序列依赖性DNA弯曲（Crothers et al. 1991）。该方案设计用于纯化的重组蛋白，但也描述了用于较粗蛋白制剂时对该方案的修改。

对于EMSA方法的理论和其他实践环节的概述，见Tried和Bromberg（1997）。理想情况下，为了得到精确的生理性的结合测量，反应中的DNA浓度应低于K_d值。由于大多数真核转录因子以$10^{-9}\sim10^{-10}$ mol/L的K_d值结合，因此DNA浓度应该低于这个值。对于一个10 μl反应，1 fmol的片段相当于约10^{-10}mol/L DNA。对50 kDa的蛋白质来说，10 μl反应中含0.5 ng，相当于10^{-9}mol/L。注意蛋白质纯化期间，由于灭活性修饰（如关键Cys的氧化），可能会丧失大量的DNA结合活性。因此，加入10^{-9}mol/L蛋白质，不能保证其活性蛋白为10^{-9} mol/L。对于纯蛋白而言，必须加入BSA或低浓度的非离子型去污剂，以阻止蛋白质与塑料试管的结合。同时必须加入载体DNA，以阻止片段上非特异性蛋白与探针的结合，从而掩盖特异性结合。加入DTT的目的是防止孵育期间蛋白质特别是Cys的氧化；加入KCl和$MgCl_2$的目的是最小化骨架磷酸的电荷排斥，并有助于稳定蛋白质和DNA结构。

时线与安排

第一步是制备含有目的位点的^{32}P末端标记DNA片段或双链寡核苷酸，下一步是灌制4%～4.5%的非变性聚丙烯酰胺凝胶，并进行预电泳以去除凝胶中的过硫酸铵。有时在凝胶电泳缓冲液中加入各种缓冲液组分，模拟反应缓冲液，但一般并没有必要这样做。之后，进行标准结合反应，并将样品上样到凝胶上。电泳4～6 h（根据DNA片段大小决定）后，将凝胶在甲醇/乙酸中固定，贴在印迹纸上，并在真空下干燥。理想情况下，反应中使用1000 cpm的标记片段。由于放射性在结合DNA和未结合DNA（两条左右的条带）间分开，因此此剂量允许在数小时或过夜后曝光出产生的放射自显影并检测出条带。

提纲

步骤 1：配制缓冲液。
步骤 2：进行电泳迁移率变动分析（EMSA）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

丙烯酰胺混合液，30%（丙烯酰胺：双丙烯酰胺=29：1）〈!〉
过硫酸铵（10%）〈!〉
结合蛋白（重组；0.5～100 ng/μl）
牛血清白蛋白（BSA；50 μg/μl）
缓冲液 D^+
　20 mmol/L HEPES-KOH（pH 7.9）〈!〉
　20%（*V/V*）甘油
　0.2 mmol/L EDTA
　0.1 mol/L KCl〈!〉
　使用前加入：
　1 mmol/L DTT〈!〉
　0.5 mmol/L PMSF（苯甲基磺酰氟）〈!〉
10×DNA 上样缓冲液
^{32}P 标记的 DNA 模板（1 fmol/μl）〈!〉
二硫苏糖醇（DTT；0.1 mol/L）
电泳缓冲液
　0.5×TBE
　1% 甘油
凝胶固定液
　20% 甲醇〈!〉
　10% 乙酸〈!〉
　在蒸馏水中配成 1L。
甘油（20%*V/V*）
$MgCl_2$（0.1 mol/L）〈!〉
Poly（dI-dC）（1 μg/μl）
5×TBE 缓冲液
N，*N*，*N'*，*N'*-四甲基乙二胺（TEMED）〈!〉

设备

印迹纸（如 Whatman 3MM，3030-917）

电泳仪
Eppendorf 管（硅化，0.5 ml）
凝胶干燥机（真空）
磷屏成像仪（可选）
保鲜膜
水浴
X 光片（可选）

程序

步骤 1：配制缓冲液

制备新鲜的缓冲液 D^+、电泳缓冲液和凝胶固定液。

步骤 2：电泳迁移率变动分析（EMSA）

1.　制备自然丙烯酰胺凝胶

a. 配制 40 ml 4.5%的自然丙烯酰胺凝胶：

溶液	体积	终浓度
丙烯酰胺混合液	6.0 ml	4.5%
5×TBE	4 ml	0.5×
20% 甘油	2 ml	1%
H_2O	28 ml	-
10% 过硫酸铵	300 μl	0.075%

b. 倒胶前，加入 30 μl TEMED。

c. 使用 1～1.5 mm 厚的隔片倒胶。

d. 在 10 mA 下预电泳凝胶 2 h。

2. 在 0.5 ml 硅化 Eppendorf 管中建立结合反应。

重组蛋白（0.5～100 ng/μl）	1.00 μl
^{32}P-DNA 模板（1 fmol/μl）	1.00 μl
poly（dI-dC）（1μg/μl）	0.20 μl
BSA（50 μg/μl）	0.25 μl
DTT（0.1 mol/L）	0.10 μl
$MgCl_2$（0.1 mol/L）	0.75 μl
缓冲液 D^+	6.70 μl
总体积	10.0 μl

必须滴定重组蛋白、DNA 模板及 poly（dI-dC）的最佳用量，使用粗提物时通常从分析中删除 $MgCl_2$。

3. 30℃（对于粗提物在室温或冰上）孵育样品 1 h。

4. 非常小心地将样品直接（即不加染料）上样到已预电泳的 4.5%的自然聚丙烯酰

胺凝胶的加样孔底。

往一个泳道上加入 5 μl 带有染料但不含反应物的 10×DNA 上样缓冲液中，作为标记，以确定产物在凝胶中已跑了多远。

5. 10 mA 电泳凝胶所需长度的时间（即以致溴酚蓝染料向下迁移至凝胶约 2/3 处）。

6. 凝胶电泳结束时，小心除去缓冲液，从电泳装置中取出凝胶，除去隔片，分开玻璃板，使凝胶附着在一块板上。

7. 将凝胶在凝胶固定液中固定 15 min。

8. 将凝胶置于两片印迹纸上，用保鲜膜覆盖在凝胶上面，在凝胶干燥机上 70℃干燥 1 h。

9. 将干燥的凝胶在放射自显影胶片或磷屏成像屏上曝光过夜。

附加说明

结合的优化一般涉及对缓冲液 pH、盐和镁离子浓度的滴定，也包括载体 DNA 和特异性标记片段之间的平衡（见 Fried 1989）。对于在短 DNA 位点上的结合测定，应使用双链寡核苷酸进行反应。当使用限制性酶切片段时，有效长度限制为 250 bp 左右（第 7 章）。

对较粗的蛋白质制剂，包括核提取物时，为了观察到特异性结合，必须进行广泛的滴定。在这种情况下，进行额外的载体 DNA 的滴定通常是必不可少的。当使用粗提物时，通常使用毫克级的提取物或更少。粗制剂中的很多蛋白质都能非特异性地与 DNA 结合，产生被迁移率变动的复合物。为确保特异性结合，使用在位点上带有突变的片段，或以特异性和非特异性寡核苷酸竞争物竞争（见第 7 章）。也可以将抗目的蛋白的抗体用于超迁移实验，以确定结合的特异性。然而，在这种情况下，应该进行仅含有抗体的反应，并与免疫前血清或以相似方法制备的一种抗体相比较。

大的蛋白质复合物（质量>1 MDa）难以进行电泳分离，且在自然聚丙烯酰胺凝胶上迁移得相当缓慢。对于分析更大的复合物，使用琼脂糖或 37.5∶1 的丙烯酰胺∶双丙烯酰胺凝胶。有些蛋白质仅在凝胶迁移实验中不能很好地结合，但与 DNase I 足迹法或另一种方法更相容（第 7 章）。这是由高 K_d值（弱结合）造成的，表明凝胶电泳期间发生广泛的解离。另外，用于凝胶的 TBE 缓冲液可能与蛋白质和其位点的结合不相容。直接在 TBE 中测试结合，有助于确定是否是这种情况。其他的低盐电泳缓冲液，包括 HEPES 和 TAE 都能代替 TBE 。有时在凝胶和缓冲液中可加入二价阳离子。在这些情况下，凝胶的导电性必须保持到最低，以防止凝胶过热；如果电泳缓冲液的缓冲能力低，应将缓冲液在电泳槽的上下槽间循环，以维持 pH。Fried 和 Bromberg（1997）曾讨论了凝胶电泳过程中影响蛋白质-DNA 复合物稳定性的一些相关参数。

疑难解答

无结合

可能原因：无结合通常说明蛋白质不够浓缩，或不能与位点结合，或缓冲液条件与

电泳期间蛋白质的结合不相容。

解决办法：试用更高浓度的蛋白质，改变缓冲液 pH 和盐浓度，使用替代电泳跑胶缓冲液，或在 4℃下或冷室中跑胶。

拖尾凝胶

可能原因：拖尾凝胶可能说明电泳期间蛋白质解离，反应物中蛋白质太多，或凝胶电泳温度过高。

解决办法：放缓凝胶电泳速度，减小蛋白质浓度，尝试在冷室中电泳，改变缓冲液条件，或加入更多的竞争性 DNA。

条带太多

可能原因：太多的条带通常是由非特异性结合造成的。

解决办法：尝试加入较少的蛋白质和较多的载体 DNA。

方案 11.6　^{32}P 末端标记 DNA 片段的准备

制备唯一^{32}P 末端标记的 DNA 片段对进行 DNA 结合实验（如 DNase I 足迹和乙基化干扰）必不可少。这里，我们将简要描述一种末端标记限制性酶切片段的方法。

时线与组织

对于带有含目的结合位点区域的质粒 DNA，该方案涉及用单一限制性内切核酸酶切割，该限制性内切核酸酶产生含磷酸的 5′ 突出端。这对于两种常见形式的片段末端标记一般都是必需的，即利用多聚核苷酸激酶磷酸化进行片段末端标记，或利用 DNA 聚合酶（如 Klenow 片段）进行“末端补平”进行片段末端标记。对于磷酸化反应，用小牛肠磷酸酶和细菌碱性磷酸酶去除磷酸，并用多聚核苷酸激酶和 [γ-^{32}P] ATP 对产生的游离 5′ -OH 进行磷酸化。这相应地产生了每个末端都标记有 γ-^{32}P 的质粒片段。为了产生唯一末端标记的 DNA 片段，对标记质粒进行热处理，使任何残留的激酶失活，然后用第二种内切核酸酶再次切割，从而释放一个短的 DNA 片段和较长的载体片段。在 5%～8%的自然聚丙烯酰胺凝胶上，将该 DNA 片段从标记载体中纯化出来。质粒 DNA 的摩尔量必须低于加入反应的 ATP 量，并且 ATP 必须具有足够高的比活性，以产生满足很多 DNA 结合实验所需的标记程度的片段。相比之下，寡核苷酸引物在 5′端通常未磷酸化，因此可以以激酶直接磷酸化，从而避开磷酸酶处理步骤。DNA 限制性酶切片段的制备和标记一般可在 1～2 天内完成。

提纲

步骤 1：质粒 DNA 的切割和标记（4 h）。

步骤 2：产生唯一末端标记的 DNA 片段（5 h）。

步骤 3：电洗脱被标记的 DNA 片段（3 h）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

过硫酸铵（10%）〈!〉
小牛肠磷酸酶（20 U/μl）
10×DNA 上样缓冲液
干冰〈!〉
乙醇（80%和 95%）〈!〉
[γ-^{32}P] ATP（水成的，10 mCi/ml，5000～7000 Ci/mmol）〈!〉
制造商提供的激酶缓冲液
酚：氯仿（1：1）〈!〉
质粒 DNA（0.5 μg/μl）
聚丙烯酰胺（40%；丙烯酰胺：甲叉丙烯酰胺＝29：1）〈!〉
多聚核苷酸激酶（10 U/μl）
蛋白酶 K（10 mg/ml）〈!〉
限制性内切核酸酶
10×限制酶缓冲液
SDS〈!〉
乙酸钠（3 mol/L）
TBE 缓冲液
TE 缓冲液
　10 mmol/L Tris（pH 7.9）
　1 mmol/L EDTA
N，*N*，*N*′，*N*′-四甲基乙二胺（TEMED）〈!〉

设备

琼脂糖制胶槽（水平）
相机闪光装置（可选）
电泳仪
Eppendorf 管（1.5 ml）
荧光染料标记（可选）
盖革计数器
加热块和（或）水浴
微量离心机
台纸
巴斯德吸管

摄影胶片（XAR-5；Kodak）
自动移液器（1 ml）
吸管（测序）
保鲜膜
剃刀刀片
闪烁计数器
辐射屏蔽
锥形管（50 ml）
透析管（平面直径 10 mm，12～14 kDa M. W. 截留）
镊子

程序

步骤 1：切割并标记质粒 DNA

一般来说，1 μg 4 kb 的质粒相当于 380 fmol DNA，切割一次后，产生 760 fmol 的末端。因此，10 μg 质粒相当 7.6 pmol DNA 末端。

1. 在 1.5 ml 微量离心管中，用能产生 5′ 磷酸突出端的酶消化 10 μg 质粒 DNA。

DNA（0.5 μg/μl）	20.0 μl
10×限制酶缓冲液 1	5.0 μl
限制酶 1	5.0 μl
小牛肠磷酸酶（20U/μl）	0.5 μl
dH_2O	19.5 μl
总体积	50.0 μl

2. 37℃孵育 2 h。

3. 加入 1 μl 10 mg/ml 的蛋白酶 K，并加入 SDS 至终浓度 0.1%，55℃孵育 15 min，消化残留的磷酸酶和限制酶。

4. 加入 50 μl TE 缓冲液和 100 μl 酚：氯仿（1∶1），轻轻涡旋，14 000g 离心 1 min。

5. 小心吸取 90 μl 水相（上层），转移至新的 1.5 ml Eppendorf 管中。

不要吸取界面部分，可以留有少许水相，这一步将有效地去除所有磷酸酶。

6. 往水相中加入 10 μl 3 mol/L 乙酸钠和 250 μl 95%乙醇。充分混匀，干冰上孵育 15 min。

7. 在微量离心机中 14 000 *g* 离心 15 min，收集 DNA 沉淀。用拉伸的玻璃巴氏吸管小心除去乙醇。

8. 加入 100 μl 80%乙醇洗涤沉淀，在微量离心机中离心 1 min。用拉伸的巴斯德吸管除去乙醇，彻底风干沉淀。

9. 将干燥的 DNA 沉淀重悬于 16 μl dH_2O 中，轻轻上下吹打数次。

10. 末端标记 DNA，混合：

DNA	16 μl
10×激酶缓冲液	3.0 μl
[γ-^{32}P] ATP (10 mCi/ml)	10.0 μl
多聚核苷酸激酶 (10U/μl)	1.0 μl
总体积	30.0 μl

采取最大限度的预防措施和屏蔽，以杜绝暴露于^{32}P，防止^{32}P泄漏。

11. 轻轻混匀，37℃孵育 15 min。

12. 65℃孵育 15min，热灭活激酶。

13. 按步骤 1：4～1：8，酚/氯仿抽提、乙醇沉淀，并洗涤 DNA 沉淀。

旁边保留一 50 ml 锥形管，储存带有放射性的拉伸吸管。

步骤 2：产生唯一标记的 DNA

1. 将 DNA 沉淀（来自步骤 1：13）重悬于 10 μl ddH_2O 中。

2. 用第二种限制酶消化 DNA。

^{32}P DNA	10.0 μl
10×限制酶缓冲液 2	5.0 μl
限制酶 2 (2 U/μl)	5.0 μl
dH_2O	30.0 μl
总体积	50.0 μl

3. 37℃消化 DNA 3 h。

4. 样品消化期间，配制 8%自然聚丙烯酰胺凝胶（用于 100 bp 范围内的片段）。

a. 在 50 ml 一次性锥形管中混合下列成分：

40% 丙烯酰胺 (29∶1)	6 ml
5×TBE 缓冲液	6 ml
dH_2O	18 ml
总体积	30 ml

b. 加入 300 μl 10% 过硫酸铵和 30 μl TEMED，充分混匀。

c. 迅速倒入凝胶板中。插入样品孔梳子，夹住凝胶顶部，将梳子夹紧。

d. 使凝胶聚合 1 h，小心拔去梳子。

e. 200 V 预电泳 1 h，去除过硫酸铵。

5. 向消化反应中加入 5μl 10×DNA 上样缓冲液。

6. 用测序吸头将该混合物上样到凝胶的加样孔中。

7. 在 1×TBE 缓冲液中电泳，直到溴酚蓝大致向下迁移至凝胶的一半。

在 8%的凝胶上，溴酚蓝与 45 bp 标记同步迁移，二甲基苯青与 160 bp 标记同步迁移。

步骤 3：电洗脱被标记的 DNA 片段

1. 非常小心地将凝胶从垂直电泳成套装置上取下。

2. 将下槽中的放射性缓冲液倒入放射性废液容器中。

3. 在水池中用水漂洗剩下的凝胶装置，用水漂洗平板。

4. 在放射屏蔽后的台纸上操作，用撬胶尺小心撬开平板，将胶留在一块平板上。

5. 用保鲜膜小心覆盖带凝胶的平板。

6. 在暗室中操作，将胶在 XAR-5 摄影胶片上曝光 3 min。

7. 用荧光染料标记凝胶周围的保鲜膜（在三处作点标记），或用相机闪光装置对凝胶闪光三次。

后一种方法在胶片上产生凝胶和平板的轮廓，前一种方法产生一系列黑点。任何一种方法都必须在后面使放射自显影和凝胶对齐，以切割相应的条带。

8. 对胶片进行显影。

9. 利用荧光标记或闪光轮廓将胶片和凝胶对齐。

10. 用剃刀刀片和镊子切出与探针对应的凝胶薄片，切通保鲜膜。

11. 将凝胶片置入一个含有 500 μl 1×TBE 的截面约为 2 in（1 in=2.54m）长的透析管。

12. 将含有凝胶片的透析管置入含有 1×TBE 的水平琼脂糖制胶槽中。

13. 100 V 电洗脱 1 h，将 DNA 从凝胶中洗脱出来。

14. 电洗脱 1 h 后，转换电极方向，电洗脱 2 min。

这将使 DNA 从管中洗脱出来，并返回到缓冲液中。

15. 用 1 ml 自动移液器从透析管中吸取含有被洗脱 DNA 片段的缓冲液。转移至 1.5 ml 微量离心管中。

16. 用等体积的酚：氯仿进行抽提。

17. 加入 1/10 体积的 3 mol/L 乙酸钠和 2 倍体积的 95%乙醇，采用如上所述的方法沉淀 DNA（步骤 1：6～步骤 1：8）。

18. 将最终的沉淀重悬于 TE 缓冲液中，浓度为 10 000 cpm/μl（以盖革计数器判断)。

19. 取 1 μl 探针点到一片 1 cm^2 的滤纸上，用闪烁计数器的 ^3H 通道测定切伦科夫放射活性。

20. 与 0.01 μCi 起始 ATP 的放射活性相比较，确定掺入程度和比活性。

一般来说，DNase I 足迹分析中使用 10 000 cpm，但 EMSA 中用量小得多。

附加说明

有时，有必要使用不产生唯一位点的酶切割限制性片段，只要 ATP 是实际 DNA 末端数目的两倍或更多，这一点就是可行的。对于 EMSA 和足迹反应来说，理想的片段长度为 50～200 个核苷酸。

许多研究者所用的另一种方法是通过 PCR 产生末端标记的片段。在这种情况下，将大约 10 ng 质粒 DNA 和各 10 pmol 的两个引物（其中一个被末端标记）一起孵育，其中两个引物分别与目的区域的两侧退火。20～30 个 PCR 循环后，产生具有高比活性的末端标记片段，并且按上文步骤 2 和步骤 3 所述在自然聚丙烯酰胺凝胶上纯化回收标记片段。PCR 比较容易，但片段的品质可能不稳定。

疑难解答

DNA 片段无标记或弱标记

可能原因：可能没有加入 DNA。

解决办法：通过测定 A_{260} 的吸光率检查 DNA 浓度，或直接对质粒进行诊断性琼脂糖凝胶，以确保 DNA 的确被加入。如果加入了 DNA，则检查激酶和缓冲液。

可能原因：陈旧的 DTT 可失活激酶。

解决办法：需要新鲜的 DTT，以维持多聚核苷酸激酶活性。

可能原因：污染性核酸的背景标记。在质粒制剂中污染 RNA，激酶也会磷酸化 RNA 末端。

解决办法：通过 CsCl-EtBr 梯度离心去除 RNA，或直接预跑 QIAGEN 柱。不要加入 RNase，它会增加末端的数量，引发更大的麻烦。

可能原因：激酶失活。

解决办法：使用新购买的激酶。

参考文献

Agresti A. and Bianchi M. E. 2003. HMGB proteins and gene expression. *Curr. Opin. Genet. Dev.* 13：170-178.

Alvarez M.，Rhodes S. J.，and Bidwell J. P. 2003. Context-dependent transcription：All politics is local. *Gene* 313：43-57.

Ausubel F. M.，Brent R. E.，Kingston E.，Moore D. D.，Seidman J. G.，Smith J. A.，and Struhl K. 1994. *Current protocols in molecular biology*. John Wiley and Sons，New York.

Balasubramanian B.，Pogozelski W. K.，and Tullius T. D. 1998. DNA strand breaking by the hydroxyl radical is governed by the accessible surface areas of the hydrogen atoms of the DNA backbone. *Proc. Natl. Acad. Sci.* 95：9738-9743.

Berger M. F.，Phiippakis A. A.，Qureshi A. M.，He F. S.，Estep P. W. 3rd，and Bulyk M. L. 2006. Compact，universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nat. Biotechnol.* 24：1429-1435.

Blackwell T. K. 1995. Selection of protein binding sites from random nucleic acid sequences. *Methods Enzymol.* 254：604-618.

Bondarenko V. A.，Liu Y. V.，Jiang Y. I.，and Studitsky V. M. 2003. Communication over a large distance：Enhancers and insulators. *Biochem. Cell Biol.* 81：241-251.

Bosisio D.，Marazzi I.，Agresti A.，Shimizu N.，Bianchi M. E.，and Natoli G. 2006. A hyper-dynamic equilibrium between promoter-bound and nucleoplasmic dimers controls NF-κB-dependent gene activity. *EMBO J.* 22；25：798-810.

Branden C. and Tooze J. 1991. *Introduction to protein structure*. Garland Publishing，New York.

Brenowitz M.，Senear D. F.，Shea M. A.，and Ackers G. K. 1986. Quantitative DNase footprint titration：A method for studying protein-DNA interactions. *Methods Enzymol.* 130：132-181.

Bulyk M. L. 2006. DNA microarray technologies for measuring protein-DNA interactions. *Curt. Opin. Biotechnol.* 17：422-430.

Burley S. K. and Roeder R. G. 1996. Biochemistry and structural biology of transcription factor IID (TFIID). *Annu. Rev. Biochem.* 65：769-799.

Carey M. 1998. The enhanceosome and transcriptional synergy. *Cell* 92: 5-8.

Carey M., Kakidani H., Leatherwood J., Mostashari F., and Ptashne M. 1989. An amino-terminal fragment of GAL4 binds DNA as a dimer. *J. Mol. Biol.* 209: 423-432.

Chen L., Glover J. N., Hogan P. G., Rao A., and Harrison S. C. 1998. Structure of the DNA-binding domains from NFAT, Fos and Jun bound specifically to DNA. *Nature* 392: 42-48.

Chinenov Y. and Kerppola T. K 2001. Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. *Oncogene* 20: 2438-2452.

Crothers D. M., Gartenberg M. R., and Shrader T. E. 1991. DNA bending in protein-DNA complexes. *Methods Enzymol.* 208: 118-146.

Crothers D. M., Drak J., Kahn J. D., and Levene S. D. 1992. DNA bending, flexibility, and helical repeat by cyclization kinetics. *Methods Enzymol.* 212: 3-29.

Dekker J. 2003. A closer look at long-range chromosomal interactions. *Trends Biochem. Sci.* 28: 277-280.

Dignam J. D., Lebovitz R. M., and Roeder R. G. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 11: 1475-1489.

Dill K. A. 1997. Additivity principles in biochemistry. *J. Biol. Chem.* 272: 701-704.

Dixon W. J., Hayes J. J., Levin J. R., Weidner M. F., Dombroski B. A., and Tullius T. D. 1991. Hydroxyl radical footprinting. *Methods Enzymol.* 208: 380-413.

Echols H. 1986. Multiple DNA-protein interactions governing high-precision DNA transactions. *Science* 233: 1050-1056.

Ellenberger T. E., Brandl C. J., Struhl K., and Harrison S. C. 1992. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: Crystal structure of the protein-DNA complex. *Cell* 71: 1223-1237.

Ellington A. D. and Szostak J. W. 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature* 346: 818-822.

Ellwood K., Huang W., Johnson R., and Carey M. 1999. Multiple layers of cooperativity regulate enhanceosome-responsive RNA polymerase II transcription complex assembly. *Mol. Cell. Biol.* 19: 2613-2623.

Fairall L., Rhodes D., and Klug A. 1986. Mapping of the sites of protection on a 5 S RNA gene by the Xenopus transcription factor III A. A model for the interaction. *J. Mol. Biol.* 192: 577-591. [Erratum: *J. Mol. Biol.* 194: 581].

Feng J. -A., Johnson R. C., and Dickerson R. E. 1994. Hin recombinase bound to DNA: The origin of specificity in major and minor groove interactions. *Science* 263: 348-355.

Fried M. G. 1989. Measurement of protein-DNA interaction parameters by electrophoresis mobility shift assay. *Electrophoresis* 10: 366-376.

Fried M. G. and Bromberg J. L. 1997. Factors that affect the stability of protein-DNA complexes during gel electrophoresis. *Electrophoresis* 18: 6-11.

Galas D. J. and Schmitz A. 1978. DNase footprinting: A simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* 5: 3157-3170.

Garvie C. W. and Wolberger C. 2001. Recognition of specific DNA sequences. *Mol. Cell* 8: 937-946.

Gralla J. D. 1985. Rapid "footprinting" on supercoiled DNA. *Proc. Natl. Acad. Sci.* 2: 3078-3081.

Grange T., Bertrand E., Espinas M. L., Fromont-Racine M., Rigaud G., Roux J., and Pictet R. 1997. In vivo footprinting of the interaction of proteins with DNA and RNA. *Methods* 11: 151-163.

Griffith J., Hochschild A., and Ptashne M. 1986. DNA loops induced by cooperative binding of λ repressor. *Nature* 322: 750-752.

Grosschedl R. 1995. Higher-order nucleoprotein complexes in transcription: Analogies with site-specific recombination. *Curr. Opin. Cell Biol.* 7: 362-370.

Grosschedl R., Giese K., and Pagel J. 1994. HMG domain proteins: Architectural elements in the assembly of nu-

cleoprotein structures. *Trends Genet*. 10: 94-100.

Ha J. H., Spolar R. S., and Record M. T., Jr. 1989. Role of the hydrophobic effect in stability of site-specific protein-DNA complexes. *J. Mol. Biol*. 209: 801-816.

Halford S. E. and Marko J. F. 2004. How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res*. 32: 3040-3052.

Hardwidge P. R., Parkhurst K. M., Parkhurst L. J., and Maher L. J., 3rd. 2003. Reflections on apparent DNA bending by charge variants of bZIP proteins. *Biopolymers* 69: 110-117.

Hayes J. J. and Tullius T. D. 1989. The missing nucleoside experiment: A new technique to study recognition of DNA by protein. *Biochemistry* 28: 9521-9527.

Hidalgo P., Ansari A. Z., Schmidt P., Hare B., Simkovich N., Farrell S., Shin E. J., Ptashne M., and Wagner G. 2001. Recruitment of the transcriptional machinery through GAL11P: Structure and interactions of the GAL4 dimerization domain. *Genes Dev*. 15: 1007-1020.

Hillisch A., Lorenz M., and Diekmann S. 2001. Recent advances in FRET: Distance determination in protein-DNA complexes. *Curr. Opin. Struct. Biol*. 11: 201-207.

Hochschild A. 1991. Detecting cooperative protein-DNA interactions and DNA loop formation by footprinting. *Methods Enzymol*. 208: 343-361.

Hochschild A. and Ptashne M. 1986. Cooperative binding of λ repressors to sites separated by integral turns of the DNA helix. *Cell* 44: 681-687.

Hollenhorst P. C., Jones D. A., and Graves B. J. 2004. Expression profiles frame the promoter specificity dilemma of the ETS family of transcription factors. *Nucleic Acids Res*. 32: 5693-5702.

Hong M., Fitzgerald M. X., Harper S., Luo C., Speicher D. W., and Marmorstein R. 2008. Structural basis for dimerization in DNA recognition by Gal4. *Structure* 16: 1019-1026.

Jantz D., Amann B. T., Gatto G. J., Jr., and Berg J. M. 2004. The design of functional DNA-binding proteins based on zinc finger domains. *Chem. Rev*. 104: 789-799.

Jen-Jacobson L. 1997. Protein-DNA recognition complexes: Conservation of structure and binding energy in the transition state. *Biopolymers* 44: 153-180.

Johnson A. D., Poteete A. R., Lauer G., Sauer R. T., Ackers G. K., and Ptashne M. 1981. λ Repressor and cro-Components of an efficient molecular switch. *Nature* 294: 217-223.

Jordan S. R. and Pabo C. O. 1988. Structure of the lambda complex at 2. 5Å resolution: Details of the repressor-operator interactions. *Science* 242: 893-899.

Kahn J. D. and Crothers D. M. 1992. Protein-induced bending and DNA cyclization. *Proc. Natl. Acad. Sci*. 89: 6343-6347.

Kslionis B. and O' Farrell P. H. 1993. A universal target sequence is bound in vitro by diverse homeodomains. *Mech. Dev*. 43: 57-70.

Kerr L. D. 1995. Electrophoretic mobility shift assay. *Methods Enzymol*. 254: 619-632.

Kim J. L., Nikolov D. B., and Burley S. K. 1993a. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* 365: 520-527.

Kim T. H. and Ren B. 2006. Genome-wide analysis of protein-DNA interactions. *Annu. Rev. Genomics Hum. Genet*. 7: 81-102.

Kim T. K. and Maniatis T. 1997. The mechanism of transcriptional synergy of an in vitro assembled interferon-β enhanceosome. *Mol. Cell* 1: 119-130.

Kim Y., Geiger J. H., Hahn S., and Sigler P. B. 1993b. Crystal Structure of a yeast TBP/TATA-box complex. *Nature* 365: 512-520.

Kissinger C. R., Liu B., Martin-Bianco E., Kornberg T. B., and Pabo C. O. 1990. Crystal structure of an engrailed homeodomain-DNA complex at 2. 8Å resolution: A framework for understanding homeodomain-DNA

interactions. *Cell* 63: 579-590.

Klemm J. D., Rould M. A., Aurora R., Herr W., and Pabo C. O. 1994. Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* 77: 21-32.

Koblan K. S., Bain D. L., Beckett D., Shea M. A., and Ackers G. K. 1992. Analysis of site-specific interaction parameters in protein-DNA complexes. *Methods Enzymol*. 210: 405-425.

Kohler J. J. and Schepartz A. 2001a. Effects of nucleic acids and polyanions on dimer formation and DNA binding by bZIP and bHLHZip transcription factors. *Bioorg. Med. Chem.* 9: 2435-2443.

Kohler J. J. and Schepartz A. 2001b. Kinetic studies of Fos · Jun · DNA complex formation: DNA binding prior to dimerization. *Biochemistry* 40: 130-142.

Koudelka G. B. 2000. Cooperativity: Action at a distance in a classic system. *Curr. Biol.* 10: R704-707.

Lee D. K., Horikoshi M., and Roeder R. G. 1991. Interaction of TFIID in the minor groove of the TATA element. *Cell* 67: 1241-1250.

Lesser D. R., Kurpiewski M. R., and Jen-Jacobson L. 1990. The energetic basis of specificity in the Eco RI endonuclease-DNA interaction. *Science* 250: 776-786.

Liang S. D., Marmorstein R., Harrison S. C., and Ptashne M. 1996. DNA sequence preferences of GAL4 and PPR1: How a subset of Zn2 Cys6 binuclear cluster proteins recognizes DNA. *Mol. Cell. Biol.* 16: 3773-3780.

Lorenz M., Hillisch A., Payet D., Buttinelli M., Travers A., and Diekmann S. 1999. DNA bending induced by high mobility group proteins studied by fluorescence resonance energy transfer. *Biochemistry* 38: 12150-12158.

Love J. J., Li X., Case D. A., Giese K., Grosschedl R., and Wright P. E. 1995. Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature* 376: 791-795.

Luscombe N. M., Laskowski R. A., and Thornton J. M. 2001. Amino acid-base interactions: A three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res*. 29: 2860-2874.

Luscombe N. M., Austin S. E., Berman H. M., and Thornton J. M. 2000. An overview of the structures of protein-DNA complexes. *Genome Biol*. 1: reviews001. 1-001. 37.

Maeder M. L., Thibodeau-Beganny S., Osiak A., Wright D. A., Anthony R. M., Eichtinger M., Jiang T., Foley J. E., Winfrey R. J., Townsend J. A., et al. 2008. Rapid "open-source" engineering of customized zinc-finger nucleases for hightly efficient gene modification. *Mol. Cell* 31: 294-301.

Manfield I. and Stockley P. G. 1994. Ethylation interference. *Methods Mol. Biol.* 30: 125-139.

Mangelsdorf D. J. and Evans R. M. 1995. The RXR heterodimers and orphan receptors. *Cell* 83: 841-850.

Marmorstein R., Carey M., Ptashne M., and Harrison S. C. 1992. DNA recognition by GAL4: Structure of a protein-DNA complex. *Nature* 356: 408-414.

Maxam A. M. and Gilbert W. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci.* 74: 560-564.

McNally J. G., Muller W. G., Walker D., Wolford R., and Hager G. L. 2000. The glucocorticoid receptor: Rapid exchange with regulatory sites in living cells. *Science* 287: 1262-1265.

Merika M. and Thanos D. 2001. Enhanceosomes. *Curr. Opin. Genet. Dev.* 11: 205-208.

Mukherjee S., Berger M. F., Jona G., Wang X. S., Muzzey D., Snyder M., Young R. A., and Bulyk M. L. 2004. Rapid analysis of the DNA-binding specificities of transcription factors wirh DNA microarrays. *Nat. Genet.* 36: 1331-1339.

Ogata K., Sato K., and Tahirov T. H. 2003. Eukaryotic transcriptional regulatory complexes: Cooperativity from near and afar. *Curr. Opin. Struct. Biol.* 13: 40-48.

Oliphant A. R., Brandl C. J., and Struhl K. 1989. Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: Analysis of yeast GCN4 protein. *Mol. Cell. Biol.* 9: 2944-2949.

Otwinowski Z., Schevitz R. W., Zhang R. G., Lawson C. L., Joachimiak A., Marmorstein R. Q., Luisi B. F., and Sigler P. B. 1988. Crystal structure of trp repressor/operator complex at atomic resolution. *Nature*

335: 321-329. [Erratum: *Nature* 335: 837].

Pabo C. O. and Sauer R. T. 1992. Transcription factors: Structural families and principles of DNA recognition. *Annu. Rev. Biochem.* 61: 1053-1095.

Pan C. Q., Finkel S. E., Cramton S. E., Feng J. A., Sigman D. S., and Johnson R. C. 1996. Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts. *J. Mol. Biol.* 264: 675-695.

Panne D. 2008. The enhanceosome. *Curr. Opin. Struct. Biol.* 18: 236-242.

Panne D., Maniatis T., and Harrison S. C. 2004. Crystal structure of ATF-2/c-Jun and IRF-3 bound to the interferon-β enhancer. *EMBO J.* 23: 4384-4393.

Panne D., Maniatis T., and Harrison S. C. 2007. An atomic model of the interferon-β enhanceosome. *Cell* 129: 1111-1123.

Parkhurst K. M., Brenowitz M., and Parkhurst L. J. 1996. Simultaneous binding and bending of promoter DNA by the TATA binding protein: Real time kinetic measurements. *Biochemistry* 35: 7459-7465.

Patikoglou G. A., Kim J. L., Sun L., Yang S. H., Kodadek T., and Burley S. K. 1999. TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev.* 13: 327-3230.

Pavletich N. P. and Pabo C. O. 1991. Zinc finger-DNA recognition: Crystal structure of a Zif268-DNA complex at 2.1Å. *Science* 252: 809-817.

Ptashne M. 1992. *A genetic switch: Phage λ and higher organisms*, 2nd ed. Cell Press and Blackwell Scientific Publications, Cambridge, Massachusetts.

Record M. T., Jr., Ha J. H., and Fisher M. A. 1991. Analysis of equilibrium and kinetic measurements to determine thermodynamic origins of stability and specificity and mechanism of formation of site-specific complexes between proteins and helical DNA. *Methods Enzymol.* 208: 291-343.

Rippe K., von Hippel P. H., and Langowski J. 1995. Action at a distance: DNA-looping and initiation of transcription. *Trends Biochem. Sci.* 20: 500-506.

Robertson G., Hirst M., Bainbridge M., Bilenky M., Zhao Y., Zeng T., Euskirchen G., Bernier B., Varhol R., Delaney A., Thiessen N., Griffith O. L., He A., Marra M., Snyder M., and Jones S. 2007. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat. Methods* 4: 651-657.

Sambrook J. and Russell D. W. 2001. *Molecular cloning: A laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schwabe J. W. and Rhodes D. 1997. Linkers made to measure. *Nat. Struct. Biol.* 4: 680-683.

Seeman N. C., Rosenberg J. M., and Rich A. 1976. Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci.* 73: 804-808.

Senear D. F., Ross J. B., and Laue T. M. 1998. Analysis of protein and DNA-mediated contributions to cooperative assembly of protein-DNA complexes. *Methods* 16: 3-20.

Shostak Y., Van Gilst M. R., Antebi A., and Yamamoto K. R. 2004. Identification of C. elegans DAF-12-binding sites, response elements, and target genes. *Genes Dev.* 18: 2529-2544.

Siebenlist U. and Gilbert W. 1980. Contacts between Escherichia coli RNA polymerase and an early promoter of phage T7. Proc. *Natl. Acad. Sci.* 77: 122-126.

Siebenlist U., Simpson R. B., and Gilbert W. 1980. E. coli RNA polymerase interacts homologously with two different promoters. *Cell* 20: 269-281.

Sigman D. S., Kuwabara M. D., Chen C. H., and Bruice T. W. 1991. Nuclease activity of 1, 10-phenanthroline-copper in study of protein-DNA interactions. Methods Enzymol. 208: 414-433.

Starr D. B. and Hawley D. K. 1991. TFIID binds in the minor groove of the TATA box. *Cell* 67: 1231-1240.

Suck D. 1994. DNA recognition by DNase I. *J. Mol. Recognit.* 7: 65-70.

Swaminathan K., Flynn P., Reece R. J., and Marmorstein R. 1997. Crystal structure of a PUT3-DNA complex

reveals a novel mechanism for DNA recognition by a protein containing a Zn2Cys6 binuclear cluster. *Nat. Struct. Biol.* 4: 751-759.

Tahirov T. H., Sato K., Ichikawa-Iwata E., Sasaki M., Inoue-Bungo T., Shiina M., Kimura K., Takata S., Fujikawa A., Morii H., et al. 2002. Mechanism of c-Myb-C/EBPβ cooperation from separated sites on a promoter. *Cell* 108: 57-70.

Thomas J. O. and Travers A. A. 2001. HMG1 and 2, and related "architectural" DNA-binding proteins. *Trends Biochem. Sci.* 26: 167-174.

Todd R. B. and Andrianopoulos A. 1997. Evolution of a fungal regulatory gene family: The Zn (II) 2Cys6 binuclear cluster DNA binding motif. *Fungal Genet. Biol.* 21: 388-405.

Travers A. A. 2004. The structural basis of DNA flexibility. *Philos. Trans. A Math. Phys. Eng. Sci.* 362: 1423-1438.

Tuerk C., MacDougal S., and Gold L. 1992. RNA pseudoknots that inhibit human immunodeficiency virus type 1 reverse transcriptase. *Proc. Natl. Acad. Sci.* 89: 6988-6992.

Tullius T. D. and Dombroski B. A. 1986. Hydroxyl radical "footprinting": High-resolution information about DNA-protein contacts and application to λ repressor and Cro protein. *Proc. Natl. Acad. Sci.* 83: 5469-5473.

Voss T. C. and Hager G. L. 2008. Visualizing chromatin dynamics in intact cells. *Biochem. Biophys. Acta* 1783: 2044-2051.

Wang J. C. and Giaever G. N. 1988. Action at a distance along a DNA. *Science* 240: 300-304.

Warner J. B., Philippakis A. A., Jaeger S. A., He F. S., Lin J., and Bulyk M. L. 2008. Systematic identification of mammalian regulatory motifs' target genes and functions. *Nat. Methods* 5: 347-353.

Watson J. D., Baker T. A., Bell S. P., Gann A., Levine M., and Losick R. 2004. *Molecular biology of the gene*, 5th ed. Pearson Education/Benjamin Cummings, San Francisco.

Weinmann A. S. and Farnham P. J. 2002. Identification of unknown target genes of human transcription factors using chromatin immunoprecipitation. *Methods* 26: 37-47.

Werner M. H. and Burley S. K. 1997. Architectural transcription factors: Proteins that remodel DNA. *Cell* 88: 733-736.

Wissmann A. and Hillen W. 1991. DNA contacts probed by modification protection and interference studies. *Methods Enzymol.* 208: 365-379.

Wolberger C. 1999. Multiprotein-DNA complexes in transcriptional regulation. *Annu. Rev. Biophys. Biomol. Struct.* 28: 29-56.

Wolberger C., Vershon A. K., Liu B., Johnson A. D., and Pabo C. O. 1991. Crystal structure of a MATα-2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell* 67: 517-528.

Yang J. and Carey J. 1995. Footprint phenotypes: Structural models of DNA-binding proteins from chemical modification analysis of DNA. *Methods Enzymol.* 259: 452-468.

Zwieb C. and Adhys S. 1994. Improved plasmid vectors for the analysis of protein-induced DNA bending. *Methods Mol. Biol.* 30: 281-294.

网络资源

http://www.biochem.ucl.ac.uk/bsm/prot_dna/prot_dna_cover.html：该网站包含通过X射线结晶学揭示的分辨率达3.0Å或更大的蛋白质-DNA复合物的结构分类，其中，蛋白质-DNA复合物通过特定DNA识别模体亚家族进行了归类。

http://www.biochem.ucl.ac.uk/bsm/sidechains/index.html：该网站提供了一个描绘在已知蛋白质结构内氨基酸侧链如何相互包装的图集。

http://www.ebi.ac.uk/thornton-srv/databases/DNA-motifs/：该网站提供了一种以已知 DNA 结合模体的一系列结构模板搜索蛋白质三维结构的方法。

http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml：人类基因组计划主页。

http://predictdnahth.rutgers.edu/：预测特定三维蛋白质结构是否含有 DNA 结合螺旋-转角-螺旋结构模体的资源。

http://the_brain.bwh.harvard.edu/：该网站提供到使用微阵列鉴定蛋白质结合位点的现行技术和软件的链接。

12　体外转录和起始前复合物组装

要点

• 可以而且应该在体外复制一个基因的特定转录吗?

• 当试图在体外复制特定转录模式时，提取物系统、模板及适当的分析方法的选择是关键。

• 可以采取几种方法来实现在体外的适当调控。

• 纯的或分馏的系统对详细的原理研究是必需的。

• 测量完整转录复合物的分析方法可用于确定激活因子如何刺激转录。

• 有许多分析方法可用于研究基因活化的个别步骤，尤其是早期步骤（如 TFIID 结合 DNA）。

• 应建立重建的转录系统以理解生化水平上激活的机制。

引言和概述

1979 年，Robert Roeder 实验室在试管中复制了一个 mRNA 编码基因的精确转录（Weil et al. 1979a）。这一里程碑式的发现使我们在分子水平上理解启动子活动的机制成为可能，这对阐明控制基因调控的通路来说是重要的一步。虽然最初的生化研究使用了病毒启动子，但现代改良技术更精确地再现了细胞基因复杂的表达模式。

在哺乳动物基因调控领域最具挑战性的问题之一是理解一个激活因子或增强体刺激转录的生化机制，目前通常的观点是，激活因子一般通过促进一个大型多组分转录复合物在核心启动子处的组装而刺激转录（综述见 Hahn 2004，第 1 章）。这种复合物由 RNA 聚合酶 II（Pol II）、通用转录因子（GTF）、共激活因子和染色质重塑组件构成。通过与原核对应物的类比，将其称为起始前复合物（PIC）。

PIC 上关键的共激活因子是调解因子（Med）和转录因子 IID（TFIID）（Lewis and Reinberg 2003；Conaway et al. 2005；Malik and Roeder 2005）（见第一章）。激活因子直接与调解因子互作使转录调控成为可能。激活因子也与 TFIID 互作。TFIID 结合 TATA 盒，尽管它与许多其他活动相关，如磷酸化、乙酰转移酶活性及泛素化（Chen and Hampsey 2002；Matangkasombut et al. 2004；Muller and Tora 2004）。激活因子、Med 和 TFIID 的联合行动，导致 GTF 的募集，从而形成 PIC（Johnson and Carey 2003）。

一旦 PIC 形成，ATP 水解使复合物异构化，并使其从所谓的“关闭复合物”转变为“开放复合物”。“关闭复合物”中起点仍然完好，而“开放复合物”中包含起点的 DNA 开始熔解（Wang et al. 1992a）。最后，Pol II 通过在 mRNA 中形成第一个磷酸二酯键而启动转录，并在产生 8-核苷酸转录本后从启动子上获释，之后沿基因延伸（Orphanides and Reinberg 2002；Woychik and Hampsey 2002）。生产性延伸需要 pTEF-b 和其他因子。在许多情况下，聚合酶重新启动转录以便生产多个 mRNA 分子。每一个步骤都可以进一步细分，并在原核系统研究的基础上，每一步骤都受到调控。

由于这些原因，生化研究都聚焦于转录过程的许多不同的步骤。在过去，大部分的重点放在 PIC 组装的早期阶段，因为这些阶段对建立基因的转录都是很重要的（Woychik and Hampsey 2002）。然而，有些研究着眼于某些激活因子对晚期步骤的影响，包括 TFIIH 的结合或活性，TFIIH 影响随后的 Pol II 负责的延伸（Kobor and Greenblatt 2002）。其他的研究支持这样的观点，即激活因子的主要效应之一是刺激重新起始（Sandaltzopoulos and Becker 1998；Yudkovsky et al. 2000）。重新起始是调控的一个逻辑点，因为延伸后，某些 GTF（如 TFIIA、TFIID、TFIIE、TFIIH 和 Med）被认为仍然保留在启动子上，而其他因子脱落并扩散掉（如 TFIIB 和 Pol II；Hahn 2004）。因此，由于复合物已经部分组装，激活因子对后随因子和（或）聚合酶分子招募的影响可能不同于初始转录复合物的形成，这种差异可能在调节转录的实际速率方面很重要。最后，体内转录发生在染色质环境中。由此可见，激活因子必须克服染色质的负面影响才可以募集通用机器（Orphanides and Reinberg 2002；Freiman and Tjian 2003；Boeger et al. 2005；Margueron et al. 2005；Mellor 2005）。当启动一个生化研究时，研究人员

必须将这些步骤的每一步都视为有可能是可调控的。

由于启动一项生化研究需要花费的费用和时间，所以必须评估为所研究的基因发展一个体外转录系统是否实际。显然，体外系统不是一个用于分析基因调控的体内方法的恰当替代。但是，如果体内研究表明这个基因是以一种有趣的方式来调控，研究人员可能会希望在体外再现该基因的转录。体外系统是推断调控的生化机制的一个重要的出发点。

如果有一个令人信服的理由要进行体外系统研究，那么当试图开发体外系统时，要考虑以下几点。

• 能从目的因子或启动子具有活性的组织和细胞系中建立体外转录系统吗？是否能获得可正确调控基因的细胞系或丰富的动物来源的适宜组织，在建立一种体外系统的决策中是一个关键的物流问题。

• 分析中用什么类型的DNA模板呢？例如，如果一个基因在体内的恰当调控需要多个控制元件的协同作用，那么模板应包括一个增强子的自然的启动子，也可以是一个人造模板，但该人造模板须含有一个功能性增强子的部件或多聚化的激活因子结合位点。对模板的选择将最终取决于研究者的需要和要解决的问题。

• 最后，基因调控需要特殊核蛋白质结构（如增强体或染色质）的重建吗？大量的研究强调了ATP依赖性染色质重建机器和组蛋白乙酰化酶、去乙酰化酶在转录激活和抑制中的作用。因此，一个基因精确的调控可能需要染色质模板的重现，以观察共激活因子和共抑制因子的效应（Kraus and Kadonaga 1998；An et al. 2004；Lavigne et al. 2004；Gamble et al. 2005；Black et al. 2006）。如果激活因子组合是必需的，可能需要大量的工作来表达和纯化它们。

许多真核基因调控的早期生化研究都是在病毒上进行的，包括SV40和腺病毒(Berk 2005)。病毒含有强组成型启动子/增强子，故使其成为富有吸引力的实验生物。而且，对其遗传学规律已有初步了解，可以将遗传学规律与生物化学、分子生物学相结合，深入了解基因调控。相比之下，多数细胞基因都以中低水平表达，并使用特殊的组织特异性激活因子和共激活因子，而且这些激活因子在提取物中的浓度可能并不高。因此，虽然病毒启动子可以作为重要的模型使用，但关键的挑战是以组织特异性方式再现重要的受调控细胞基因的转录。本章中，我们讨论了达到这一目的所能采用的策略，使研究人员可以避开分析中会遇到的一些技术性问题和实验假象。技术细节可参考原始文献，这里我们尝试概括主要的决策过程。

如上文所述，若研究者计划对负责转录调控的因子做更高级的研究，体外转录分析会非常有价值。发展一种体外转录分析，有如下要求。

• 可获得足够大量细胞的细胞来源，以制备高浓度蛋白质提取物。

• 目的基因转录起始位点的相关知识。

• 测定特定启动子活性的分析方法。

• 理想情况下，当分析复杂的启动子时，获得一种或多种重组调控蛋白会简化分析过程，这是因为提取物中这些因子浓度很低，会给高级研究带来问题。

激活机制的研究涉及选择合适的模板。基于模式模板和自然启动子的系统在该领域

被广泛使用。如果一个特定的模板已用于体外转录，那么它也应该用于机制研究。通常情况下，模式系统是由一个单一激活因子蛋白和承载多聚化位点的报告基因组成的。在这些系统中所做的研究可以提供一个特定的激活因子是如何在机制水平上起作用的一般生化认识。然而，有数项实验已造成这样的看法，即核蛋白结构（如增强体）是在启动子或增强子上形成的（Ogata et al. 2003；Li et al. 2004；Boeger et al. 2005）。正是增强体最终与染色质重塑机器和通用转录机器相互作用。因此，当启动一项分析时，一个必须要问的关键问题是，一个单个激活因子的机制是否能提供重要的信息，或者是否有必要解决来自具有其所有固有复杂性的自然启动子的调控问题。在一定程度上，这个问题涉及两个主要争论点：①一个可将转录和结合研究连接在一起的体外转录系统的可得性；②考虑一下，关于该系统，已经从体内研究或以前的生化分析中对其有多少了解。

活化转录的体外成功再造提供了理解该机制细节方面的重要机会。关于体外激活因子靶向的具体步骤知之甚少。在不同的激活因子如何与通用转录机器，尤其是与调解因子那样的共激活因子如何沟通方面，有一个很大的认识的空白。造成这一空白的第一个原因是，该领域的许多实验室已经转向激活因子对染色质重塑效应的研究。第二个原因是，许多实验室都不愿意研究 PIC 组装所需的大型共激活因子复合物（即 TFIID 和 Med），因为分离和表征它们有困难。但是，有了带有表位标记复合物的细胞系，以及简便的新的分析方法（如固定化模板），纯化和表征成为比过去更简单的工作。

一旦选定系统并有试剂，则应问下列问题：一个激活因子是否组装一个有功能的转录复合物，它是否在复合物形成后才起作用，或两者兼而有之？如果一个激活因子刺激复合物组装，那么哪些步骤会受到影响？是否激活因子互作并招募主要共激活因子复合物 TFIID 和 Med？染色质对转录复合物的组装的影响是什么？在这一章中，我们总结了怎样使用一些方法来产生当前的哺乳动物转录和基因激活的生化见解。转录复合物组装的某些方面进行了非常深入的研究，而关于其他方面的研究很有限。在描述这些实验时，我们在某种程度上采取按时间的先后顺序排列的方法。这并不是说早期的方法已经过时，实际上，许多早期的方法现在仍然被广泛使用。我们首先描述功能研究，其中信息读出是转录。但是，主要的重点是测定转录过程中机制性步骤的分析方法。

虽然我们讨论基础和激活两种转录，但基础系统显示在缺乏调控性因子的情况下因子建造复合物的顺序。此外，从这样的分析获得的信息可以阐明重要的调控检查点。不过，重要的是要考虑基础转录是由一个包括 Pol II、TATA 盒结合蛋白（TBP）、TFIIA、TFIIB、TFIIE、TFIIF 和 TFIIH 等因子的小集合完成的。许多额外的多肽，包括 TBP 相关因子（TAF）、Med 及染色质修饰和重塑酶等，对激活因子在体内的功能发挥是必不可少的。并不完全清楚这些对激活必要的额外的蛋白质是否改变转录过程的基本机制方面，或简单地改变转录复合物形成的数量。

应该很明显，各种方法都有一定的局限性，而实验设计无疑会影响研究结果。在一种混合物中，各种因子的浓度和所使用的方法学可给反应强加一个特定的限速步骤，而在其他情况下（即在体内）这种反应可能并不是极限的。但是，在没有遗传学的情况下，往往难以评价生化研究的有效性。当可获得时，应使用其中所研究的因子已经被破坏的细胞系来研究体内作用（Stevens et al. 2002）。其中，RNA 干扰的使用已被证明是

一个用于确认在体内共激活因子重要性的流行方法（Pavri et al. 2006）。此外，通用转录因子间从酵母到人类的高度保守性使得酵母系统在破解真核基因激活机制中极其有用(Smale and Kadonaga 2003；Hahn 2004；Blazek et al. 2005)。因此，在酵母中的实验可以帮助评估实验结果在哺乳动物系统中的有效性。

实验策略

提取物的制备

细胞的选择

体外转录分析可以利用组织或细胞提取物进行。提取物制备过程中需要大量细胞，一些组织（如鼠肝）比较廉价，可以提供足量的细胞。由于细胞类型异质、蛋白酶解和细胞裂解困难等，固体组织提取物问题较多。考虑到转录需要大量的特异性转录因子和通用转录因子，这些问题可能会阻碍高效体外分析方法的发展。因此，制备用于体外转录分析的提取物时，常用转化细胞，且我们主要侧重于培养细胞。

很难为提取物制备推荐一个确切的细胞数，因为这取决于要解决的问题所需要的不同的细胞量。一般情况下，很难从少于 10^9 个细胞准备提取物，因为蛋白质提取物在制备过程中必须保持浓缩。使用小量细胞会导致浓度较低的提取物，因为较小的缓冲液体积和较少的细胞沉淀物比较大的体积可导致更大的损失。此外，某些操作，包括 Dounce 匀浆、试管转移、透析，用较小体积更难进行。不过，也有一些方法，它们可以使用替代裂解技术并尽量减少操作次数。这些技术已成功地用于制备小量提取物（如方案 12.1）。大多数在体外分析中取得成功的实验室，都使用从 10^9～10^{10}个或更多的细胞中制备的提取物。

为了从组织培养细胞中准备提取物，细胞的健康指数增长很重要。例如，HeLa 细胞在每升转瓶培养物中生长到 5×10^8 细胞。不像很多其他细胞系，HeLa 细胞可在二氧化碳培养箱外培养，并且生长培养基可高压灭菌，还能添加便宜的牛或马血清。通常情况下，如果细胞在 37℃温控箱或标准暖房内孵育，以 4 L 平底 Florence 培养瓶生长 2 L HeLa 细胞，则可从这些培养物中制备 2～3 ml 提取物。对于长期使用细胞提取物的实验室，建议建立每周可生长 32～50 L 细胞的设施。如果将细胞放置在可保持恒定缓冲和营养补充环境的特殊生长箱或细胞培养发酵罐中，较小体积的细胞可生长到较高的密度。但是这些设备价格昂贵，且需要经常性保养维修。

不像 HeLa 细胞及某些淋巴细胞，许多贴壁培养的细胞（如成纤维细胞）不适合于旋转培养，在 10 cm 平板上只能产生 10^6～10^7个细胞。即使在这种规模的高端水平，为了产出 1～2 L HeLa 细胞，也需要培养 100 个这样的平板。此外，可使用滚瓶培养扩大培养规模，特别是在旋转培养下生长状态不良的淋巴细胞。然而，与适于旋转培养的细胞相比，这些细胞的大规模培养仍有较大的难度。从不能在旋转培养条件下生长的细胞中制备提取物耗时且很困难。由于这种原因，研究人员应该估量体外研究结果是否值得投入这么大的人力、物力和财力。为估算这一后勤问题，可假定一个研究者想从标准贴壁细胞系的 100 个细胞平板着手，根据培养细胞 24～28 h 的倍增时间特征，需要花费 2

周或更长的时间，才能处理完 100 个平板。提取物制备之前收集的细胞沉淀物体积大约为 1 ml，利用常规核抽提程序能产生 2 ml 提取物，而典型的体外转录实验大约使用 20 μl（100～200 μg）提取物。因此，2 ml 提取物平均能用于 100 个反应。一次典型的实验涉及 20 个反应，因此 2 周工作所获得的提取物可用于 5 次体外转录实验。

在转录研究领域中，将基因的组织特异性表达在体外重现是最难的方法之一，需要投入大量的时间和资源。由于细胞特异性激活因子或共激活因子的差异很大，只有在基因具有正常活性的组织或细胞系中才能实现正确的体外调控。如果所有或许多转录激活因子能以重组形式获得，并能以外源形式加入，则比较容易进行。这种情况下，异源组织（如 HeLa 细胞）可用作通用转录因子容易而丰富的来源，之后，向这些异源组织补充外源激活因子。

对于那些对酵母体外转录系统感兴趣的研究者，如果具有酵母中等规模生长的设施（http://www.fhcrc.org/science/labs/hahn/methods/biochem_meth/Pol II_trans.html），制备提取物相当容易。酵母提取物活性高，并已广泛使用，取得了关于共激活因子活动的许多基本观点（Ranish et al. 1999；Yudkovsky et al. 2000）。果蝇的提取物也被用来研究体外转录，并已经有制备活性胚胎或细胞系提取物的好方法（Parker and Topol 1984；Kamakaka and Kadonaga 1994）。

提取物制备方法

Dignam 和 Roeder 程序（Dignam et al. 1983a）是从组织培养细胞中制备核提取物的方法（方案 12.1），已经在多个细胞系，包括 HeLa 细胞、B 细胞（BJAB/Namalwa）和前列腺癌细胞系（LNCaP）中获得成功。该方法涉及收集细胞沉淀，并将其重悬在过量的低渗缓冲液中。这一处理使细胞膨胀，膨胀的细胞可以通过标准手动 Dounce 匀浆器进行破碎。小心地匀浆会使细胞破碎，而较小的细胞核会保持完整，通过低速离心将这些细胞核与残留细胞和细胞质碎片分离。细胞核可以重悬在中盐缓冲液（0.38～0.42 mol/L NaCl）中，这种缓冲液将转录成分从细胞核中抽提出来。已经表明所产生的提取物含有组蛋白，这可能归因于稍微有限的核裂解，但在 Dignam 提取物中重建染色质的努力尚未获得成功。

值得注意的是，用于 Dignam/Roeder 方案的 KCl 浓度（0.42 mol/L）是根据经验确定的腺病毒启动子转录的最佳浓度（Dignam et al. 1983a）。换言之，用此 KCl 浓度所得提取物中既含转录所需的正作用的转录因子，也具有干扰转录的负作用转录因子，但两者之间恰当平衡。可以推断，低盐浓度抽提正作用转录因子的效率较低，而较高盐浓度可能导致降低启动子活性的其他抑制性蛋白质的提取（如组蛋白 H1 就是在 0.6 mol/L KCl 条件下释放的）。请记住这一点，通过使用不同的 KCl 浓度制备提取物，从而为一个新的启动子优化提取物是明智的。

将蛋白质从细胞核中抽提出来后，可以直接透析，也可以通过硫酸铵沉淀。对于许多细胞系和基因而言，将提取物进行硫酸铵沉淀可以增加体外转录活性。硫酸铵沉淀通常作为早期步骤用于蛋白质纯化过程，以达到较小的纯化效果，因为在所有相关因子发生沉淀所需的浓度下，有些污染性蛋白质不沉淀（Scopes 1994）。硫酸铵沉淀是浓缩提取物最有用的方法。未沉淀的提取物，特别是来自小细胞体积的未沉淀提取物，蛋白质

浓度往往是 2～5 mg/ml。通过有效地沉淀蛋白质，并使其重新溶于小体积的溶剂中，能将提取物浓缩至 10～20 mg/ml。此外，可用 Microcon 蛋白浓缩器浓缩提取物。这种浓缩处理能使典型的体外转录反应物中加入更多的蛋白质。提取物经硫酸铵沉淀并重悬后，可对提取物进行透析，以减少盐浓度。

虽然 Dignam 法已成功地广泛用于从组织培养细胞中制备提取物，但尚未应用于固体组织，这是因为固体组织难以产生单细胞悬液。单细胞悬液对于有效膨胀细胞并用于手工 Dounce 匀浆来说是必需的。但是有许多替代方法可用于制备固体组织提取物（Schewe et al. 2003；Banerjee et al. 2005；Rashi-Elkeles et al. 2006）。一个特别成功的组织提取方法是使用大鼠肝提取物纯化有活性的 GTF 和共激活因子（Conaway et al. 1996）。一个典型的方法如下所述。首先用解剖刀和剪刀在冷的等渗缓冲液中将组织剪碎，之后将组织与低渗缓冲液孵育。然后，将切碎并膨胀的组织用机械 Dounce 匀浆器匀浆。如果组织足够细致地被切碎，手工 Dounce 匀浆也是可以的。将匀浆后的组织铺到密度梯度蔗糖上，从细胞碎片中沉淀并分离出细胞核。可以用 Dignam 法或裂解细胞核的变通方法抽提这些细胞核，并沉淀染色质，抽提转录组分（Gorski et al. 1986）。

抽提固体组织提取物的主要原则是，目的基因在该组织中表达，并且该组织是基因特异性调控蛋白的理想来源。但是，如果固体组织提取物本身并不能支持基因活化，则可以往 HeLa 提取物中补充来源于固体组织提取物的纯化的调控因子或色谱馏分。HeLa 提取物是 GTF 和遍在细胞激活因子（如 SP1）的良好来源。对提取物补充外源组织特异性激活因子足够在体外再现基因调控的某些层次。假如这些激活因子在 HeLa 提取物中丰度不高，那么可以采用对提取物补充外源组织特异性激活因子的方法，该方法在多种激活因子中都已证明是成功的。GAL4-VP16 是作为外源加入的一种激活因子，因为在 HeLa 提取物中没有识别 GAL4 位点的蛋白质（Carey et al. 1990a，b）。的确，如果内源激活因子存在，并且它对再现激活因子活性是必要的，就可能不得不通过免疫沉淀或亲和层析将其去除。例如，Kim 和 Maniatis（1997）通过 DNA 亲和层析从哺乳动物提取物中去除了 β-干扰素（IFN-β）增强子结合蛋白。如果后面的这些方法失败，对提取物进行分馏有时能将激活因子从通用转录因子中分离出来。

如果利用核粗提物或硫酸铵沉淀的提取物没有得到特异的启动子活性，可以尝试其他提取物。例如，尝试制备全细胞提取物（Manley et al. 1980，1983），这种提取物中的其他蛋白质可能对某些启动子具有更高的活性。另一种有用的策略是利用肝素-琼脂糖树脂进行粗分级。利用这种树脂，所有的基础转录因子和许多 DNA 结合蛋白都可以用 0.2～0.5 mol/L KCl 进行洗脱，而其他许多抑制转录的蛋白质不与柱结合（Dynan and Tjian 1983）。要进行肝素-琼脂糖层析，应将粗提物或硫酸铵沉淀的提取物首先透析成柱-载样缓冲液。可能的话，测定电导率，确保盐浓度已经在透析中降到足够低。然后，透析后的提取物可以上肝素-琼脂糖柱（至少 1～2 ml）。原则上，每毫升树脂可承载大于 10 mg 的蛋白质，并且最高达 40 mg 蛋白质。虽然 40 mg 的蛋白质量超过了柱的容量，但这可以确保洗脱出的蛋白质浓度相当大。但是，当对少量提取物进行操作时，这种做法是不现实的。应该权衡过柱过程中提取物量的中小程度的潜在损失和获得浓度较高的提取物的好处。采用小（1～2 ml）层析柱（Bio-Rad Econo Colum）时，为

获得有效的蛋白质结合，可调整柱压为2～3倍重力进行层析。采用大层析柱时，蛋白质应该以1～4倍柱体积/h的标准速度过柱。结合后，先用载样缓冲液洗涤，之后用含有0.1 mol/L或0.2 mol/L KCl的载样缓冲液洗涤。蛋白质用含有0.5 mol/L KCl的载样缓冲液洗脱，然后将洗脱物进行透析，并检验其活性。

转录分析

一般考虑

要为一个新基因建立体外转录系统，首先要确定提取物能利用强启动子，这一证明步骤通常在优化提取物、缓冲液条件和启动子的初始阶段完成。作为一个标准模板，可以使用那些仅依赖于基础转录因子或遍在上游激活因子的启动子。腺病毒主要晚期启动子（AdMLP）是一个指导高水平体外转录的常用启动子。AdMLP含有TATA和Inr元件，能在缺乏上游激活因子的情况下指导高水平体外转录，AdMLP也含有USF（上游刺激因子）蛋白和其他可增强转录的激活因子的结合位点。如果这一启动子在目的细胞提取物中不起作用，就需要购买或制备HeLa提取物。HeLa提取物可用于检修体外转录分析的基本环节。

首先要利用对照启动子对模板和提取物的浓度进行滴定，因为两者都会影响启动子活性。对于对照启动子和目的启动子，提取物和模板的最佳浓度可能相似，但不会相同，因此对每种启动子都应该进行滴定。如果用对照启动子检测到了强活性，但用实验启动子未检测到强活性，最好系统地改变实验条件，以获得启动子特异的活性。后面的一节中我们将讨论一些常用的可选方法。

有多种分析方法可用于检测启动子活性，这些分析方法可分为两类。第一类，转录反应利用非标记的NTP进行，产生的mRNA利用引物延伸法、S1核酸酶或RNase保护分析法进行检测。这些方法涵盖在第3章中。引物延伸法是最有效的传统分析方法，该方法对从体外转录反应中得到的RNA产物特别有效。如果使用这种方法，使用一个不与污染粗提物的内源性细胞mRNA相结合的引物非常重要。定量反转录聚合酶链反应（RT-PCR）技术是另一种选择，但就我们所知，它尚未被广泛用于体外转录，也不一定能提供起始位点信息。第二类分析方法涉及用^{32}P标记NTP进行对转录本的直接标记。第二类分析方法中有两种类型的分析方法，即失控转录分析和无G序列盒分析（专题12.1）。在选择方法的过程中有两个重大的物流方面的考虑：①处理转录产物所需的时间（即从试管到放射自显影或磷屏成像仪）；②产物的比活性（即显示出信号所需的时间）。从这一方面讲，引物延伸法较慢，但一般比直接标记更敏感；直接标记比引物延伸分析速度更快，通常涉及向反应混合物中加入一种α-^{32}P标记的NTP。但是，放射性核苷酸必须用未标记的相同NTP广泛稀释，使总浓度接近Pol II的K_m值，否则转录效率低。

此外，在哺乳动物提取物中进行的反应一般每反应需要1～10 μCi放射性，即5～10 μmol/L α-^{32}P标记的核苷酸，以及200～500 μmol/L的其他核苷酸。由于用于一般实验的放射性量很高，产物的处理必须非常谨慎，以免污染实验室表面和设备。

专题 12.1 测定体外转录的方法

引物延伸

如第 3 章中所述，引物延伸是检测转录的常见方法（图 12.1）。在对起始位点进行体内定位的方法中，引物延伸是适于体外转录的方法之一。在该方法中，分离 mRNA，并与位于起始位点下游 50～150 个核苷酸的^{32}P 末端标记引物杂交（方案 3.1）。引物通过反转录酶延伸，被标记的 cDNA 产物在 8%～10%聚丙烯酰胺/尿素凝胶上分离。一般情况下，延伸反应体系中加入 0.1 pmol 引物就已足够，因为 0.1 pmol 引物大大超过了 mRNA 的量，能够促进大量的引物杂交和延伸。因为对引物定量标记在理论上是可行的，因此，原则上引物延伸要比无 G 转录分析或失控转录分析敏感很多倍（因为每个 mRNA 都可以被^{32}P 标记引物延伸）。但实际上，引物的标记效率、杂交效率及反转录酶的延伸都不是定量的。因此，引物延伸的敏感性要比宣称的低，但无论如何，它比[α-^{32}P] NTP 内部标记要更有效（见下文）。

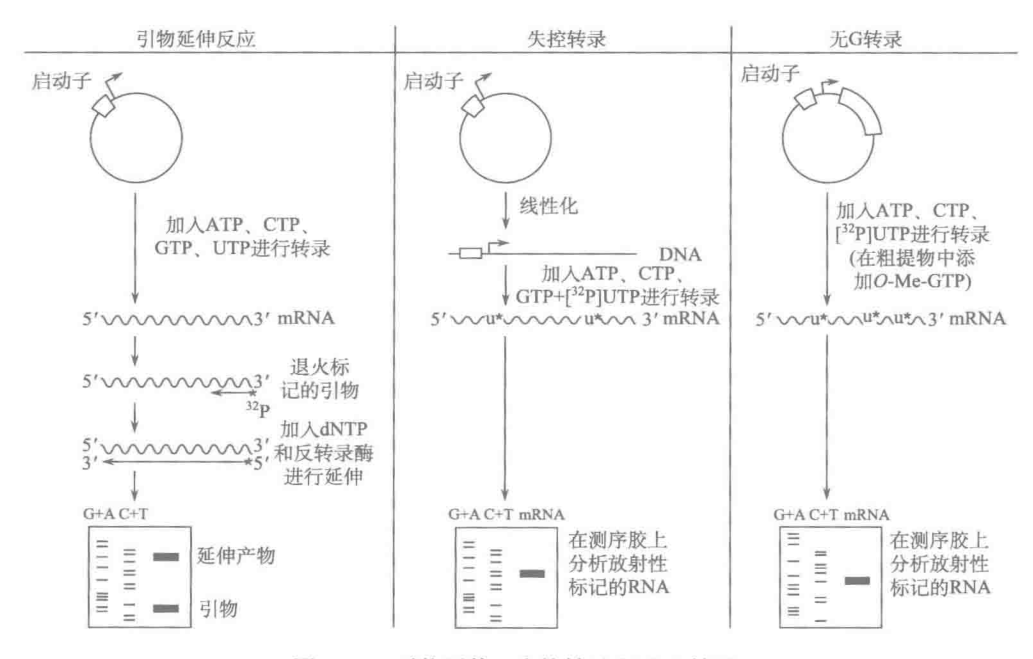


图 12.1 引物延伸、失控转录和无 G 转录。

引物延伸的一个优点是该方法仅检测短片段 RNA，因此提取物中的内源性 RNase 和样品操作过程中引起的对 RNA 的小程度切割并不妨碍分析。这一点在体外特别重要，因为在产生一个体外产物所需要的操作中，难以抑制 RNase。但是，如第 4 章中所述，研究人员必须警惕转录本的二级结构，因为二级结构会限制反转录酶的延伸效率。引物延伸反应中加入放线菌素 D 能防止伪倒扑合成（snapback synthesis）。伪倒扑合成可产生不代表天然起始位点的引物延伸产物。另外，引物延伸反应具有较窄的最佳 pH

范围，特别是利用禽源反转录酶的反应，必须在仔细缓冲的反应中进行。引物设计和实验建立中的其他考虑，见第3章及其附属方案。

利用从天然组织（该处基因正常表达）制备的提取物，在天然模板上进行转录反应时，一个潜在的问题是，提取物经常受到目的基因转录的mRNA的污染。当尝试测定一个体外转录信号时，这些污染物可能会带来难以克服的背景问题。由于这一原因，经常将异源序列或引物标签克隆于启动子下游90 bp，并代替与天然下游序列互补的引物。一个常用的方法是将启动子亚克隆到载体多克隆位点，并利用旁侧序列引物进行引物延伸。

与无G序列盒转录分析和失控分析相比，体外引物延伸的主要缺点是进行体外转录反应和引物延伸反应所花费的时间。在无G序列盒转录分析和失控分析中，内部标记的RNA可以直接加样到凝胶上。但是，当研究一种新的启动子时，由于其增高的灵敏性及在体外鉴定起始位点的能力，引物延伸显示了超过其他方法的明显优势。

失控分析

要进行失控分析（图12.1），一个典型的模板是位于启动子下游300 bp或更下游的线性化片段，在α-^{32}P放射性标记的NTP存在的条件下，检测转录。通常，反应物中含有三种浓度为200 μmol/L的未标记核苷酸和一种浓度为10～20 μmol/L的标记核苷酸。由于放射性核素的高比活性，以及能加入每个反应体系的放射性量的限制，少量被标记的核苷酸可用较低浓度、未被标记的同一种核苷酸稀释。举例来说，含有20 μmol/L未标记UTP的反应中加入1～10 μCi［α-^{32}P］UTP。由于RNA Pol II的NTP K_m的原因，UTP浓度接近观察到有效转录所需的最小量。在核苷酸存在的条件下，Pol II将起始并延伸至DNA模板的末端，产生放射性标记的mRNA。

虽然失控转录较为简单，但转录本的标记效率不高。例如，在含有20 μmol/L UTP的50 μl反应体系中，加入10 μCi［α-^{32}P］UTP（鉴于约9000 Ci/mmol的理论最大值），对于每纳摩尔未标记UTP来说，可转化1 pmol标记UTP。因此，1000个U残基中只有1个被放射性标记。鉴于长度为200个核苷酸的转录本中平均含有50个U，每个反应中，20个转录本中只有1个转录本含有1个放射性U。将未标记UTP浓度降至5 μmol/L或10 μmol/L，或提高标记核苷酸的量会改善标记效率。但是，比活性从来不能被增加到超过α-^{32}P标记核苷酸制剂的比活性，因此能加入反应的放射性量有实际的限度。反应一旦结束，反应产物就可以在聚丙烯酰胺/尿素凝胶上分级分离。40 cm的测序凝胶是首选，较短的凝胶跑胶时间较短。除非比较多个起始位点或者如果有污染性RNA在附近迁移，实际的分辨率未必重要。

靠近启动子的模板的线性化能抑制转录，因为在提取物中存在大量与DNA末端相结合的蛋白质。因此，为谨慎起见（如果没有必要），以一种仍然可使研究者观察到RNA产物的方式，在距离启动子尽量远的部位使模板线性化。

使用放射性标记NTP时一个潜在的问题是，提取物中的背景转录相当高。在很多质粒模板DNA上，都可发生非特异性RNA Pol III转录。而且Pol I、PolII、PolIII都能进行一种所谓的“末端到末端”转录反应。在该反应中，聚合酶结合模板的末端，并向内转录。可以通过加入延伸抑制剂α-鹅膏蕈碱来区分不同的产物。在哺乳动物系统中，2 μg/ml鹅膏蕈碱能抑制Pol II，200 μg/ml鹅膏蕈碱能抑制Pol III，Pol I对所有

浓度的鹅膏蕈碱都不敏感，但能被放线菌素 D 抑制（放线菌素 D 插入双链 DNA，并且是很多 DNA-酶反应的通用抑制剂）。最后，哺乳动物提取物富含一些酶，这些酶能将核苷酸加到提取物中污染性 RNA 的 3′端。因为最丰富的污染物是 5S RNA 和 tRNA，它们趋向于被^{32}P 末端标记，并产生凝胶上 65～120 碱基的条带。α-鹅膏蕈碱抑制的缺乏，以及在无 DNA 模板的情况下标记 RNA 的出现，是鉴定这类反应产物的常见方式。

在一般情况下，失控分析比使用无 G 序列盒的分析方法效率低，这主要是因为无 G 反应使用环形超螺旋 DNA，这种 DNA 模板比线性 DNA 更好。但是，失控分析确实有一个较大的优点：有时，在启动子区插入一个无 G 序列盒会破坏转录起始位点下游的启动子序列。如果发生了这种情况，明智的做法就是采用失控转录分析。

无 G 序列盒

无 G 序列盒（图 12.1）是检测转录最简单最快速的方法之一（Sawadogo and Roeder 1985b）。无 G 序列盒通常是一个非编码链上缺乏 G 残基的 360 bp 序列。产生的信使 RNA 不含 G。在一般情况下，通过将该序列亚克隆到转录起始位点，可将 360 个核苷酸的无 G 序列盒引入模板。一般可用 PCR 来完成克隆。

一般情况下，将模板与 ATP、CTP、[α-^{32}P] UTP 及未标记的 UTP 一起在提取物中孵育 1 h。为减少背景转录，不加 GTP。因反应体系中不含 GTP，原理上，转录应该在无 G 序列盒的末端终止，并且小分子非特异性 RNA 的合成应最小化。产物纯化后，将混合物上样到 6%聚丙烯酰胺/尿素脉凝胶上电泳。如前详述，1000 个 U 残基中只有一个会被放射性标记。

在粗提取物中，即使充分透析后，仍存在足够大量的污染性 GTP，它们会使聚合酶通读无 G 序列盒。在这种情况下，观察不到明显的无 G 转录本，但可以看到背景转录。一种解决办法是用 RNase T1 处理反应，该酶在 G 处切割产物。以此消除背景转录本，并将无 G 序列盒 mRNA 修剪成在凝胶上可见的独特的 360 核苷酸 RNA。但是，还有从无 G 序列盒上游起始并通读无 G 序列盒的随机转录本。因此，用 RNase 处理不能区分正确起始 mRNA 和随机起始的通读无 G 序列盒的 mRNA。为了减少这种通读转录，可往反应体系中加入过量的 GTP 类似物 3′-*O*-Me-GTP。3′-*O*-Me-GTP 缺少 3′-OH，可作为链终止类似物。在 3′-*O*-Me-GTP 存在的条件下，当聚合酶往延伸的转录本中掺入 G 时会熄火。加入 3′-*O*-Me-GTP 试剂能显著增强在 GTP 污染的提取物中进行的反应的特异性。

有很多大小各异的无 G 序列盒，较小型的无 G 序列盒可以和内对照模板相融合，用以归一化不同反应的效率。最后，应该评价用无 G 序列盒取代下游序列的效应，以确保没有意外消除关键性调控序列。最近，许多实验室的研究都表明，紧挨起始位点的下游序列可促进核心启动子的活性（见第 1 章）。

模板的选择

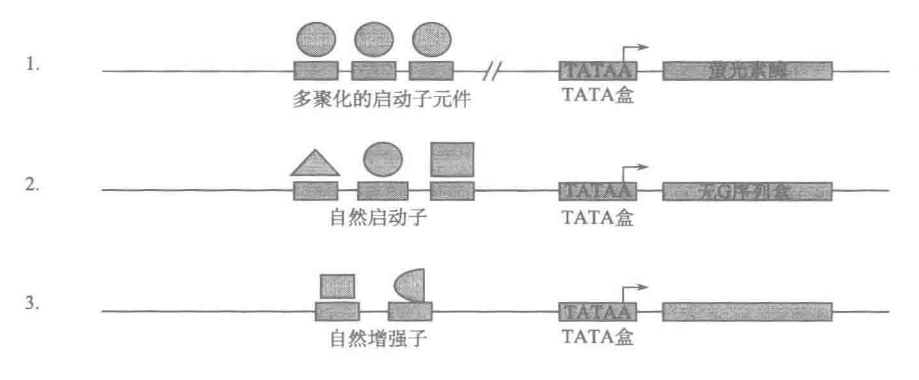
对体外转录分析选择恰当的模板是任何研究的一个重要方面（图 12.2A）。多数利用裸 DNA 模板的体外系统都不能忠实地重现距离转录起始位点超过数百个碱基对的启动子调控元件的作用。但是，在有些情况下，特殊的染色质模板让远距离效应得到观察。染色质模板也是一些受调控启动子研究的首选。

模板的选择取决于分析的目的。如果目的是理解在自然上下文序列中多种调控蛋白之间的协作效应（cooperative effect）和协同效应（synergistic effect），就应该使用天然启动子（图 12.2A）。因为远距离调控元件在体外可能不能被精确调控，这些远距离调控元件有时可通过构建它们在其中被定位于核心 TATA 盒附近的模板加以研究（例如，通过删除介于其间的 DNA 或在紧挨清楚表征的核心启动子的上游安置一个增强子）。如果目的只是简单地分析单个调控蛋白的特征，就可构建一个含有多聚化串联拷贝应答元件的人工模板（图 12.2B）。在所有情况下，最好使用经过转染分析的启动子。转染分析结果可以和体外转录结果相比较，以确定生化系统是否能忠实地再现通常在体

A. 自然模板



B. 使用各种不同调控元件和报告基因的模式模板



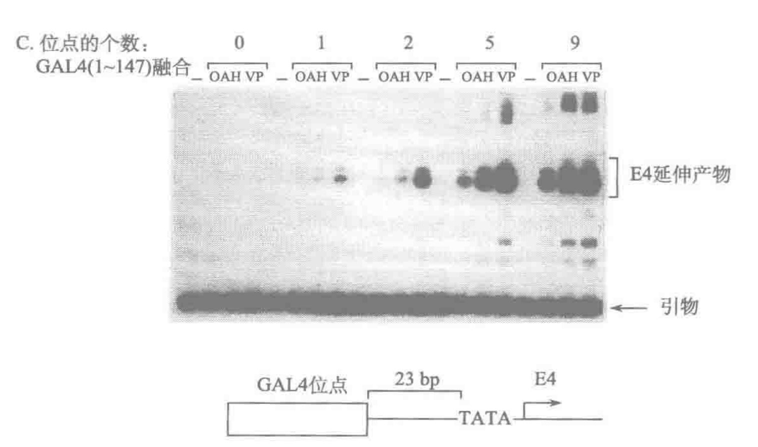


图 12.2　用于体外转录的模板（A、B 基于 Carey et al. 1990a 的数据。C 经许可转载自 Carey et al. 1990b，Magazines Ltd.）。

内看到的基因表达的恰当水平。图 12.2C 展示了一个典型分析的结果，在该分析中，将含有不同数目多聚化 GAL4 位点的人工模板与含有强度逐渐增强的激活结构域 GAL4（1～147）、GAL4-AH 和 GAL-VP16 在 HeLa 提取物中一起孵育。结果显示了 HeLa 系统的动力学范围，其中转录水平对位点数目和激活结构域强度作出应答。而且，转染分析中观察到类似的诱导模式。

通常，在优化启动子体外表达时，难以区分由上游启动子元件驱动的转录（活化转录）和单独由核心启动子驱动的转录（基础转录）。随着模板浓度的提高，基础转录胜过活化转录，因为激活因子被稀释，并分布在不同的模板分子中。而且，高浓度的模板也会促进基础因子的结合，由此避免对激活因子介导的基础因子募集的需要。在这些条件下，核心启动子元件就能独自驱动表达。为了说明所观察到的转录受到调控，应将含有目的上游激活因子结合位点的模板与只含有核心启动子的模板相比较，以说明条件对受调控转录是适合的。

在其中添加外源性激活因子的情况下，应在有无激活因子的情况下进行 DNA 模板的滴定。激活因子的最佳用量可以通过在重组蛋白中的滴定及每种模板浓度下转录的测定而确定。注意，高 DNA 模板浓度会导致需要加入非常高浓度的激活因子，以完全覆盖模板位点。在这些条件下，通过转录压制（transcription squelching），激活因子浓度可造成抑制（Carcy et al. 1990a）。在这种情况下，降低模板应允许利用较低浓度的激活因子。

不能观察到裸 DNA 上的激活作用，并不一定意味着激活因子未直接刺激 PIC 的组装。提取物或转录系统中基础水平高的转录会掩盖激活因子的效应。当转录复合物组装与染色质竞争时，激活因子的刺激效应可以更明显。几种哺乳动物启动子和增强子的活性已经在染色质模板上得以重现，包括从 TCR-α、HIV-1 和 β-珠蛋白基因上的激活（Sheridan et al. 1995，1997；Barton and Emerson 1996；Mayall et al. 1997）。许多人工增强子已经用于体外转录，包括带有 HIV-1 基础启动子的 Notch 应答元件（Fryer et al. 2002）、带有 AdMLP 基础启动子的 STAT 应答元件（Zakharova et al. 2003），以及 β-联蛋白应答元件（这是一个 LEF1 结合位点，但在缺乏 β-联蛋白时 LEF-1 结合非常弱；Tutter et al. 2001）。

重要的是要注意，经选择用于体外转录分析的模板同样可以用于结合研究。借助于生物素化引物进行的启动子区 PCR 扩增，可将启动子片段结合到链亲和素包被的珠子上，以用于蛋白质结合研究（见固定化模板方法）。类似的方法允许对一个小的、明确的 DNA 片段进行电泳迁移率变动分析（EMSA）或光交联研究。下面我们将讨论这些方法。

染色质系统

有几种无细胞系统会将质粒组装成染色质模板，包括在合适条件下与 DNA 模板孵育的果蝇胚胎提取物，以及纯化的组蛋白的混合物或核小体（Becker et al. 1994；Pazin et al. 1994，1996；Bulger et al. 1995；Barton and Emerson 1996；Steger et al. 1997；Loyola et al. 2001；Fyodorov and Kadonaga 2002，2003；Carrozza et al. 2003；Loyola and Reinberg 2003；Lusser et al. 2005）。这些系统纯度各异，有些含有大量染色质修饰

活性，如 ATP 依赖性 SWI2/SNF2 重塑复合物、带有 SWI2 同源物（如 ISWI）的几个相关复合物或组蛋白乙酰化酶。为了避免往模板中加入几个染色质相关蛋白，使用纯组蛋白八聚体的连续盐稀释液的方法非常有用，尽管核小体间距不一定是周期性阵列（Carrozza et al. 2003）。这些模板的组装和它们的体外转录分析被简要地涵盖在第13章。

条件的优化

进行体外转录研究的初始步骤是优化分析条件，以便在体外重现体内或转染研究中观察到的表达模式。当优化反应条件时，需要考虑的几个参数包括：提取物和模板的浓度、缓冲液条件和对特殊补充物的需要。

提取物优化通常涉及在 25～50 μl 反应混合液中对 20～250 μg 的粗提物进行滴定（通过 Lowry 或 Bradford 分析法测定）。当滴定提取物时，模板的量应保持恒定：基于质粒的环形模板用 50～500 ng DNA，线性片段模板用较少的 DNA。相反，当滴定模板时，提取物的量应保持恒定。在多数情况下，提取物富含非特异性 DNA 结合蛋白，加入不同数量的载体 DNA（如质粒载体），对于防止转录抑制是必要的。

缓冲液和孵育条件也很关键。一般情况下，在含有 50～100 mmol/L 单价盐（NaCl、KCl、乙酸铵、乙酸钾）、5～10 mmol/L $MgCl_2$ 或 $MgOAc_2$，以及 20 mmol/L HEPES 或 50 mmol/L Tris（pH7.9）的缓冲液中，体外转录反应的活性最好。核苷酸可按 0.2～1 mmol/L 的浓度范围加入。标准时程表明，反应在 30℃孵育远远长于 1 h 也不会合成产物，想必是因为提取物中某些组分因共价修饰、热不稳定性或内源蛋白酶而失活。对于每种提取物和模板，都应优化盐、镁离子、缓冲液条件、时程和温度。也应测试体积排阻试剂（如聚乙二醇或聚乙烯醇）或特殊补充物（如亚精胺）的影响，因为这些物质常常能增强核酸代谢有关的生化反应。

有些研究人员发现在系统中加入参考质粒是有利的。这样做可以显示出第二种转录本，该转录本可作为产物回收的对照，或者用以确定一套缓冲液条件是否对多种模板都是最佳的。但是，一种效应的再现性通常是更好的策略，因为加入外源模板能引起对照模板和测试模板之间的竞争。无论如何，如果不得不包含内对照，则必须确定对照模板不受测试模板存在的影响，反之亦然。

粗提物中通常含有大量能扣押产物的 RNA 结合蛋白。因此，转录反应结束后，在酚抽提和乙醇沉淀之前，产物通常于 55℃在含有 0.2% SDS 的缓冲液中用蛋白酶 K 处理 15 min，从而使这些蛋白质降解。酚抽提后最好再进行一次酚氯仿抽提，以彻底去除蛋白质。单独的酚/氯仿抽提并不能有效地去除蛋白质，并且给产物的后续处理带来严重的困难（如凝胶上的弥散带型，或乙醇沉淀后出现不溶性沉淀等）。

分级分离的系统

Matsui 等（1980）首次证明了 HeLa 细胞提取物可通过磷酸纤维素柱层析分为不同的馏分。方法是将提取物加到含有 0.1 mol/L KCl 的磷酸纤维素柱上，然后用盐浓度逐步增加的缓冲液洗脱。Matsui 及其同事分离得到几种单独存在时无活性，而加在一起则重现特异性转录的组分。这些组分称为 TFIIA（0.1 mol/L KCl 洗脱物）、TFIIB

(0.35 mol/L KCl 洗脱物)、TFIIC (0.5 mol/L KCl 洗脱物) 和 TFIID (1 mol/L KCl 洗脱物)。TFII 名字源自于 Pol II 的转录因子。TFIIC 随后被表明是一个蛋白质 (多聚 ADP-核糖聚合酶),它通过结合 DNA 模板上的缺刻 (nick) 抑制非特异性转录,从而为特异性转录增加可用的 RNA 聚合酶量。因此,该蛋白质不是真正的通用转录因子 (GTF),尽管它可能影响染色质的转录。重要的是要注意,其他实验室也设计了分级分离 GTF 的有效的方法,并且较早的文献中这些因子通常不总是使用 TFII 命名术语 (Conaway et al. 1996)。

Reinberg 和 Roeder (1987) 随后利用其他层析树脂,如 DEAE-纤维素和肝素-琼脂糖凝胶,将这些组分进一步分离,得到其他几种因子。除了 Pol II 之外,已发现至少有 6 种不同的通用因子 (TFIIA、TFIIB、TFIID、TFIIE、TFIIF 和 TFIIH) 为转录所需要。这些因子和来自上游刺激活性 (USA) 共激活因子馏分的共激活因子相结合 (Orphanideset al. 1996),足以支持体外基础转录和活化转录。同时,USA 馏分包含大量的不同活性,其中包括 Med。许多编码不同因子的多肽已被纯化至均质状态,同时它们的基因已被克隆 (见专题 12.2)。有许多来源可获得表达表位标记因子的表达载体和细胞系 (表 12.1)。

表 12.1 纯化人转录因子表达系统的部分名单

转录因子		表达系统	参考文献
TFIIA	1. 大亚基	细菌	Dejong and Roeder (1993)
	2. 小亚基		Dejong et al. (1995)
TFIIB		细菌	Ha et al. (1991)
TFIID		人细胞:	
		HeLa (FLAG-标签 TBP)	Chiang et al. (1993)
		HeLa (HA-标签 TBP)	Zhou et al. (1992)
		HeLa (FLAG-TAF4)	Guermah et al. (2001)
		F9 (FLAG-TAF4)	Perletti et al. (2001)
		3T3 (FLAG-TAF4b)	Geles et al. (2006)
TFIIE	1. 大亚基	细菌	Chalut et al. (1994)
	2. 小亚基		
TFIIF	1. 大亚基	细菌	Wang et al. (1993)
	2. 小亚基		
PC4 (USA 组分)		细菌	Ge and Roeder (1994)
TFIIH 复合物		人细胞系:	
		HeLa 和成纤维细胞包含 HA-	Winkler et al. (1998)
		标记 XPB 和 HeLa (FLAG-ERCC3)	Malik and Roeder (2003)
TFIIH 亚基		杆状病毒表达重组亚基;	
		可重建成复合物	Tirode et al. (1999)
p300		杆状病毒	Schiltz et al. (1999)
PCAF		杆状病毒	Yang et al. (1996)
SWI/SNF 复合物		人细胞系:	
		Jurkat (HA-标记 BRG1)	Wang et al. (1996)
		HeLa (FLAG-标记 Ini1)	Schnitzler et al. (1998)

续表

转录因子	表达系统	参考文献
SWI/SNF 亚基	杆状病毒表达重组亚基； 可重建成复合物	Phelan et al. (1999)
人调解因子复合物	人细胞系：	
	HeLa 细胞	
	FLAG 标记 SRB10 (CDK8)	Gu et al. (1999)
	FLAG-标记 Nut2	Malik and Roeder (2000)
	其他 FLAG-标记亚基	Sato et al. (2004)

许多细菌和杆状病毒的构体以 His6 标签蛋白的方式供货。His6 标签蛋白可以用 Ni-NTA 亲和层析纯化（QIAGEN）或作为 FLAG 标签蛋白纯化。HA 单克隆抗体树脂和肽可从 Covance 获得。FLAG 单克隆的 M1 和 M2 单克隆抗体树脂和肽可从 Covance 和 Sigma 获得。对于许多转录因子的多克隆抗体可从 Santa Cruz 生物技术公司获得。Pol II 单克隆抗体可从 Covance 和 QED 获得。cDNA 可从引用文献和 ATCC 获得。

因子缺失的系统

如果研究目的是测定特异性激活因子对于单一通用因子的影响，那么有多种快速而简单的方法从 HeLa 提取物中大量去除假定靶因子。这些因子缺失的系统往往是高度活跃的，因为它们含有许多实现激活转录所必需的共激活因子。在每一种情况下，选择性地缺失一种因子，从而维持剩余因子的适当平衡，这对赋予核提取物高转录活性来说是必要的。例如，VP16 激活结构域与 TFIID 和 Med 都发生相互作用（Johnson et al. 2002；Johnson and Carey 2003）。

Med 可使用两种方法去除，包括：抗体耗竭和激活因子亲和柱。例如，谷胱甘肽-*S*-转移酶（GST）与甲状腺受体配体结合结构域的融合蛋白在 HeLa 提取物中和 Med 紧密互作。在配体 T3 存在的情况下，将提取物通过 GST-甲状腺激素受体层析柱，能显著消耗提取物中的 Med。耗竭后的提取物不再支持 GAL4-VP16 激活的转录，除非回加纯的 Med。当与结合分析相结合时，这个实验可以确立 Med 结合和转录激活之间的直接效应。

同样，提取物可耗尽 TFIID，以便可以往该提取物中重新补加 TFIID。将 HeLa 细胞提取物于 45℃热处理 15 min，足以使 TFIID 彻底失活，而对其他因子仅产生中等程度的影响（Nakajima et al. 1988）。这种系统不需要特殊试剂或层析设备，因此是最便于建立的系统。第二种方法是将 0.5 mol/L KCl 预平衡的核提取物缓慢流经磷酸纤维素柱，流出的溶液不含 TFIID，可以补充同一柱上 0.85 mol/L KCl 洗脱的 TFIID 组分。要注意的是，这一组分还含有 USA 共激活因子（Ge et al. 1996）。在某些情况下，TFIID 组分可用 DE52 和（或）肝素-琼脂糖凝胶柱进一步分级分离，以纯化和浓缩 TFIID。另外，可以通过免疫亲和层析耗尽提取物中的 TFIID，这一过程通常涉及将提取物与包含 TBP 和几个 TAFII 亚基（如 TAF1 和 TAF9）的抗体混合物共孵育（Oelgeschlager et al. 1998；Johnson et al. 2002）。

可以通过使核提取物缓慢流经等体积（不超过二倍体积）的镍-次氮基三乙酸（Ni-NTA）-琼脂糖凝胶柱去除 TFIIA。TFIIA 的大亚基可以与 Ni 紧密地结合。产生的提取物可以补充重组 TFIIA，或天然 TFIIA 可以用咪唑洗脱，用于和 TFIIA 缺失提取物

互补（Ozer et al. 1994）。

也可以利用亲和纯化的高效价 TFIIB 抗体实现 TFIIB 的去除，这些抗体可以从多种学术或商业来源（如 Santa Cruz Biotechnology Inc.）获得。将含有与蛋白质 A-琼脂糖凝胶相偶联的 TFIIB 抗体树脂与 HeLa 提取物一起孵育，就能从提取物中大量去除 TFIIB。这种提取物可以作为研究 TFIIB 活性的系统。以相似的方式可去除 TFIIH、Pol II、TFIIE、TFIIF 和多种共激活因子。

高度分级分离的系统

可在体外产生由所有单独的、很多呈重组形式的通用因子组成的系统，这些系统支持中低水平的转录活化（Ge et al. 1996；Maldonado et al. 1996；Tantin et al. 1996）。这种系统的建立涉及从 *E. coli* 中纯化重组人 TFIIA、TFIIB、TFIIE、TFIIF 和 PC4，从 HeLa 提取物中免疫纯化 TFIID、TFIIH、Pol II 和 Med（Thompson et al. 1990；Zhou et al. 1992；Malik and Roeder 2003）。随着我们对活化过程认识的深入，这些系统的纯度和效率会不断增加。

专题 12.2　纯化的转录因子

Malik 和 Roeder（2003）发表了纯化每个单个通用转录因子、Pol II、Med 和 PC4 的方法。

TFIID

TFIID 的 TBP 亚基是第一个被克隆的通用因子（Hahn et al. 1989）。虽然 TBP 从酵母中以单一多肽的形式被纯化，但是在哺乳动物细胞中 TBP 与 TBP 相关因子（TAF_{II}）稳定结合。在 HeLa 提取物中，有 10 个或更多个 TAF_{II}和 TBP 抗体发生免疫共沉淀。虽然 TFIID 可以通过传统方法进行纯化，但是表达血凝素（HA）或 FLAG 标记 TBP 的 HeLa 细胞系已被独立地制备。抗原表位被置于 TBP 的氨基端或氨基端附近，在这些部位抗原表位对 TFIID 的生化活性无抑制作用（Zhou et al. 1992；Chiang et al. 1993）。对重组的 HeLa 细胞系进行旋转培养，从这些细胞中制备提取物，并使其流经磷酸纤维素柱。将 0.5～0.85 mol/L 洗脱物（图 12.3）和含有 HA 或 FLAG 抗体偶联蛋白质 A-琼脂糖的免疫亲和树脂一起孵育。将非结合型蛋白质用高盐缓冲液洗去后，用 1 mg/ml HA 或 FLAG 肽竞争 TFIID，将 TFIID 从树脂中洗脱。其他实验室还创建了表达亲和标记 TAF 亚基的细胞系，包括表达 FLAG-TAF4b 的 3T3 细胞（Geles et al. 2006）、表达 FLAG-TAF4 的 F9 细胞（Perletti et al. 2001），以及表达 FLAG-TAF4 的 HeLa 细胞（Guermah et al. 2001）。这种方式产生的 TFIID 被高度纯化，可以用于体外转录、DNase I 足迹分析、EMSA 和交联分析。45℃ 热处理 HeLa 提取物 15 min，可以选择性地灭活 TFIID，从而提供一个便利的测定 TFIID 活性的系统（Nakajima et al. 1988）。但是，这只是一个可用于测定 TFIID 的有点简单的系统，而且其他活性也被不同程度地灭活。冷冻电镜（电子显微镜）的研究揭示了与 TFIID 功能相关的结构上的认识（Grob et al. 2006）。

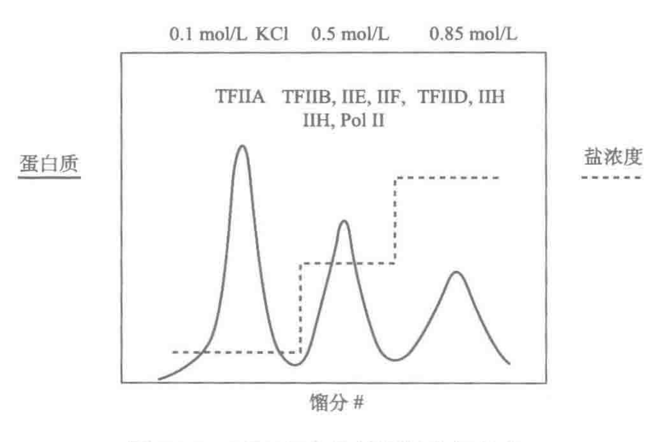


图 12.3　GTF 组分的柱层析分级分离。

TFIIB

TFIIB 是继 TBP 之后，第二个被克隆的通用因子。TFIIB 是一种含 316 个氨基酸的 33 kDa 蛋白质。TFIIB 是由 Ha 等（1991）纯化到同质，并对其进行微测序分析，以及从文库中克隆得到其 cDNA。已将其 cDNA 插入含有 FLAG 或 6×-His 标签的 T7 表达载体，并能在 *E. coli* 中易于过表达，且能纯化成同质的 TFIIB。通常要对细菌提取物先进行聚乙烯亚胺和硫酸铵处理，以分别去除核酸和杂蛋白。再用抗 FLAG 抗体或镍-琼脂糖微珠纯化 TFIIB。TFIIB 能强有力地稳定 TBP 与 TATA 寡聚核苷酸在天然聚丙烯酰胺凝胶或镁-琼脂糖 EMSA 中的结合。在这些系统中也能进行一些激活因子的募集研究，因为一些激活因子直接与 TFIIB 接触，并且募集就是一种功能结果。已经解析了单独的 TFIIB（Nikolov et al. 1995）及 TFIIB 与 RNAPol II 复合物（Bushnell et al. 2004）的晶体结构。

TFIIA

TFIIA 首先是从酿酒酵母（*Saccharomyce cerevisia*）中分离并克隆的（Ranish et al. 1992），它由两个亚基组成：TOA1 和 TOA2。遗传学研究和四分体分析表明，两个亚基对酵母生存都是必不可少的。利用酵母基因序列，从人和果蝇中克隆了 TFIIA（DeJong and Roeder 1993；Ma et al. 1993；Yokomori et al. 1993，1994；Bernstein et al. 1994；Ozer et al. 1994；Sun et al. 1994；Deiong et al. 1995）。人 TFIIA 是一种由大小分别为 37kDa、19kDa、13kDa 的三个亚基（α、β、γ）组成的复合物，并与 TFIID 的一个亚群相结合，三亚基 TFIIA 蛋白中的两个最大亚基来源于与酵母 TOA1 基因同源的一个更大的前体。该前体经蛋白酶解，产生两个成熟的亚基。亚基可以单独标记并在细菌中表达，然后亚基按化学剂量结合获得 TFIIA 复合物（Yu et al. 2001）。

TFIIA-TBP-TATA 三元复合物的足迹研究（Lieberman 1994；Lieberman and Berk 1994）和晶体结构分析表明，TFIIA 能通过一个亚基与 TBP 直接接触、另一个亚基与 DNA 接触来稳定 TFIID 和 TATA 盒的结合。其他研究表明，在生化分析中 α-β 前体能与 γ 亚基一起发挥作用。在 TBP 和 TFIID 结合分析和体外转录分析中，重组

TFIIA 具有完整的功能。

TFIIF

TFIIF 含有两个亚基，分别称为 RAP30 和 RAP74，这样命名是因为这两个亚基最初是在亲和层析实验中以 RNA 聚合酶相关蛋白（RAP）的形式分离得到的。两个亚基基因都已经被克隆（Sopta et al. 1989；Aso et al. 1992；Finkelstein et al. 1992）。可以以细菌表达载体的形式得到这些基因，并且可以在 *E. coli* 中大量表达和纯化蛋白质（Wang et al. 1993，1994）。可以混合重组亚基，它们保留了 TFIIF 的所有已知功能属性。亚基以异源四聚体的形式相缔合，通过酵母中的冷冻电镜研究已经确定了 RNA Pol II/TFIIF 的结构（Chung et al. 2003）。

TFIIE

TFIIE 含有大小为 34 kDa 和 56 kDa 的两个亚基。与 TFIIF 一样，这些亚基以异源四聚体的形式相缔合。这两个亚基都已经被克隆（Ohkuma et al. 1990；Inostroza et al. 1991；Peterson et al. 1991），且可从 T7 表达载体上进行表达。在体外转录分析中两个亚基都具有功能（Chalut et al. 1994）。使用冷冻电镜已经确定了人 TFIIE 的结构（Jawhari et al. 2006）。

TFIIH

因为具有大量不同的亚基、许多生化活性及在 DNA 修复和转录中的双重作用，TFIIH 成为研究得最多的通用因子之一。这种蛋白质已在哺乳动物细胞和酵母中（称作“因子 b”）进行了广泛的研究。TFIIH 至少含有 8 个亚基，具有数种生化活性（综述见 Egly 2001；Reese 2003；Zurita and Merino 2003）。高度纯化的 TFIIH 具有 ATP 依赖性 DNA 解旋酶活性。TFIIH 还具有激酶活性，在 ATP 存在的条件下，会使 Pol II 的羧基端结构域磷酸化。磷酸化导致 SDS 凝胶中最大亚基的迁移率降低，使 Pol II 由 IIA 形式转变到了 IIO 形式，这种作用不需要 DNA 或转录机器其他组分参与。

可以利用抗多种不同亚基的抗体以传统的层析法对 TFIIH 进行纯化，并通过免疫印迹法监测层析图谱。尽管普遍相信，存在很多 TFIIH 的层析复合物，但这些复合物具有不同子集的亚基。这些不同的复合物可能使 TFIIH 在转录起始和 DNA 修复中具有不同的作用。使用抗 TFIIH 亚基的抗体通过免疫亲和层析也已经纯化了 TFIIH（Winkler et al. 1998），还曾使用抗 FLAG 表位的抗体从 FLAG-p62 稳定 HeLa 细胞系（Kershnar et al. 1998）和 FLAG-ERCC3 稳定 HeLa 细胞（Malik and Roeder 2003）中纯化了 TFIIH。此外，已经利用杆状病毒中表达的亚基产生了完整的 TFIIH（Tirode et al. 1999；Jawhari et al. 2002b）。

Pol II

哺乳动物 RNA Pol II 含有 12 个亚基，这些亚基已经全部被克隆，但尚未实现由重组亚基重建得到 Pol II。可以利用抗羧基端结构域（CTD）的单克隆抗体，通过常规层析法或免疫亲和层析法，将 Pol II 纯化成同质分子（Thompson et al. 1990；Thompson and Burgess 1996）。也可以从 FLAG-RPB9（Kershnar et al. 1998；Malik and Roeder 2003）或 FLAG-HisRPB3 稳定的 HeLa 细胞系中纯化 Pol II（Hasegawa et al. 2003）。基本程序是，首先制备标准 Dignam 核提取物，随后进行 DEAE 纤维素分级分离和免疫

亲和或金属离子亲和纯化。最初的结构研究只包括来自酵母的核心 Pol II，它缺少一个 Rpb4/Rpb7 异二聚体（Cramer et al. 2000，2001）。从那以后，已经揭示了完整的酵母 RNA Pol II 结构（Armache et al. 2003；Bushnell and Kornberg 2003）。Pol II 与核酸的结构支持一个模型，在该模型中 DNA 进入一个由酶的“上下颌”形成的通道，并继续进入环绕含镁离子的活性位点的深的裂缝。活性位点造成 DNA 形成 90°的尖锐弯曲，从而引导核酸上游末端沿“内壁”到达酶的出口（Gnatt et al. 2001）。

基本起始前复合物的形成

转录复合物的概念最初源自研究人员观察到纯化的真核核心聚合酶自身不能识别并转录启动子，而哺乳动物提取物则能识别并转录启动子。后来对提取物的分级分离和表征揭示了 GTF 和共激活因子的存在及其身份（表 12.2～表 12.4）。该领域的关键问题是各种因子如何组装成复合物，以及哪些步骤受到激活因子和抑制因子的调控。在这一节中，我们讲述一些选取的研究，它们以解密各种因子如何组装成复合物为其研究目的。一些比较老的研究仍然是与此有关的，因为它们代表了阐明机制的较为简单的方法。然而，应该知道关于转录的基本机制已经知之甚多。虽然描述的某些分析方法可适用于由激活因子组装的 PIC，但如果要获悉独特的或新的信息，应只使用特殊分析方法（如光交联的使用）。例如，目前很少对具有独特激活因子、TFIID 和 Med 的活化转录复合物进行交联研究。此外，某些激活因子可能对各种共激活因子的结合有相当独特的影响。

表 12.2　通用转录因子，包括单个的亚基、功能和参考文献

GTF	亚基(克隆文献)		功能	参考文献
	酵母	哺乳动物		
TFIIA	TOA1(Ranish et al. 1992)	α/β(Dejong and Roeder 1993；Ma et al. 1993)	激活；TBP 相互作用；SAGA 相互作用	Conaway et al. (1990)；Weideman et al. (1997)；Ranish et al. (1999)；Kraemer et al. (2001)；Warfield et al. (2004)
	TOA2(Ranish et al. 1992)	γ(Ozer et al. 1994；DeJong et al. 1995)	激活；人类中的 αβ 旁系同源；TAC 复合物	Mitsiou and Stunnenberg (2000)；Ozer et al. (2000)；Upadhyaya et al. (2002)
TFIIB	SUA7(Pinto et al. 1992；Tschochner et al. 1992)	TFIIB(Ha et al. 1991；Tsuboi et al. 1992)	TBP 相互作用；Pol II 相互作用；起点选择；锌带结构域；B-指结构域；PIC 形成；启动子清除	Conaway et al. (1987)；Pinto et al. (1992)；Lagrange et al. (1998)；Bushnell et al. (2004)；Chen and Hahn(2004)；Tubon et al. (2004)；Pal et al. (2005)；Baek et al. (2006)
TFIIE	TFA1(Feaver et al. 1994)	α(Ohkuma et al. 1991；Peterson et al. 1991)	招募 TFIIH；启动子熔解/清除；锌指结构域；翼状螺旋-转角-螺旋结构域	Conaway et al. (1991)；Maxon et al. (1994)；Ohkuma and Roeder(1994)；Kim et al. (2000)；Meinhart et al. (2003)；Watanabe et al. (2003)；Okuda et al. 2004

续表

GTF	亚基(克隆文献)		功能	参考文献
	酵母	哺乳动物		
	TFA2(Feaver et al. 1994)	β(Peterson et al. 1991; Sumimoto et al. 1991)	DNA 结合;激活	Okamoto et al. (1998);Forget et al. (2004);Lin and Gralla (2005)
TFIIF	TFG1(Henry et al. 1992)	RAP74(Aso et al. 1992; Finkelstein et al. 1992)	Pol II 启动子打靶;延伸;TFIIB 相互作用;激活因子相互作用	J. W. Conaway and Conaway (1989);Yan et al. (1999);Reid et al. (2002);Chung et al. (2003);Ghazy et al. (2004);Zhang et al. (2005)
	TFG2(Henry et al. 1992)	RAP30(Sopta et al. 1989; Horikoshi et al. 1991)	动摇非特异性 Pol II 相互作用;起点选择;疾病	J. W. Conaway and Conaway (1989);Killeen and Greenblatt (1992);Wei et al. (2001);Ghazy et al. (2004)
	TFG3,TAFT4 (Henry et al. 1992)	AF9	常见于 TFIIF、TFIID 和 SWI/SNF;应激应答	Kimura et al. (2002);Kimura and Ishihama(2004)
TFIIH	SSL2,RAD25 (Park et al 1992)	XPB,ERCC3 (Weeda et al. 1990)	启动子熔解/清除;转录和切出修复需要的 ATP 酶/解旋酶(3′—5′),解旋酶超家族 2	R. C. Conaway and Conaway (1989);Bradsher et al. (2000);Coin et al. (2004);Lin et al. (2005);Fan et al. (2006);Lame and Egly(2006)
	RAD3(Naumovski and Friedberg 1982)	XPD,ERCC2 (Weber et al. 1990)	切出修复需要的 ATP 酶/解旋酶(5′—3′);激活;连接核心 CAK 复合物;调控 CAK;与 XPB 形成环结构	Schultz et al. (2000);Chen et al. (2003);Dubaele et al. (2003);Riedl et al. (2003);Theron et al. (2005)
	TFB1(Gileadi et al. 1992)	p62(Fischer et al. 1992)	切出修复;激活;血小板-白细胞 C 激酶底物同源结构域	Gervais et al. (2004);Liu et al. (2005);Di Lello et al. (2006)
	TFB2(Feaver et al. 1997)	p52(Marinoni et al. 1997)	切出修复;转录	Marinoni et al. (1997);Jawhari et al. (2002a)
	SSL1(Yoon et al. 1992)	hSSL1 (Humbert et al. 1994)	切出修复;锌结合;RING 结构域;泛素连接酶	Seroz et al. (2000);Takagi et al. (2005)
	CC1	cyclin H(Fisher and Morgan 1994;Makela et al. 1994)	带 Kin28 的 TFIIK;受 CDK8 调控;激活	Akoulitchev et al. (2000);Keogh et al. (2002);Bour et al. (2005)
	TFB4(Feaver et al. 1997)	p34(Humbert et al. 1994)	切出;转录;与 SSL1 相互作用	Feaver et al. 1999);Araujo et al. (2000)

续表

GTF	亚基(克隆文献)		功能	参考文献
	酵母	哺乳动物		
	TFB3(Feaver et al. 1997)	Matl(Yee et al. 1995)	锌 RING 指;TFIIK 相互作用;细胞周期;调控	Gervais et al. (2001);Rossi et al. (2001);Jona et al. (2002);Korsisaari et al. (2002);Lee et al. (2005)
	KIN28(Simop et al. 1986)	Cdk7(Fisher and Morgan 1994; Levedakou et al. 1994)	带 CCT 的 TFIIK;细胞周期;转录;热激应答	Larochelle et al. (1998);Schwartz et al. (2003);Fisher(2005)

表 12.3　TFIID 相关因子

因子	酵母	哺乳动物	功能	参考文献
TAF1	Taf145/130	TAF250	激酶,乙酰转移酶;基因表达;细胞周期;激活因子特异性;溴域	Lee et al. (2000);Solow et al. (2001);Wassarman and Sauer (2001);Maile et al. (2004);Irvin and Pugh(2006)
TAF2	Taf150	TAF150	DNA 结合;细胞周期	Verrijzer et al. (1994);Chalkley and Verrijzer(1999);Martin et al. (1999)
TAF3	Taf47	TAF140	细胞活力;组蛋白折叠	Gangloff et al. (2001)
TAF4	Taf48	TAF130/135	组蛋白折叠;形成八聚体样结构;激活因子特异性	Jacq et al. (1994;Mengus et al. (1997;Selleck et al. (2001)
TAF4b		TAF105	B 细胞和 T 细胞发育;NF-κB 介导的应答;精子发生;H2A-样	Yamit-Hezi and Dikstein (1998;Silkov et al. (2002);Falender et al. (2005)
TAF5	Taf90	TAF100	PIC 组装;细胞周期	Apone et al. (1996);Dubrovskaya et al. (1996)
TAF5L		PAF65β	精子发生;TAF5-样;PCAF 亚基	Ogryzko et al. (1998);Hiller et al. (2001)
TAF6	Taf60	TAF80	DNA 识别;H4-样;形成八聚体样结构;凋亡;发育	Burke and Kadonaga(1997);Aoyagi and Wassarman(2001);Bell et al. (2001);Selleck et al. (2001)
TAF6L		PAF65α	TAF6-样;PCAF 亚基;与 TAF9 相互作用	Ogryzko et al. (1998)
TAF7	Taf67	TAF55	激活因子特异性;调控 TAF1	Chiang and Roeder(1995);Munz et al. (2003);Gegonne et al. (2006)
TAF7L		TAF2Q	精子发生	Pointud et al. (2003)

续表

因子	酵母	哺乳动物	功能	参考文献
TAF8	Taf65	BAB71460	分化;细胞活力;组蛋白折叠	Hernandez-Hernandez and Ferrus (2001);Guermah et al. (2003)
TAF9	Taf17	TAF32/31	DNA 识别;H3-样;形成八聚体样结构;激活因子特异性	Burke and Kadonaga(1997);Uesugi et al. (1997);Selleck et al. (2001)
TAF9b		TAF31L	细胞生长;抑制;与 TAF9 相似	Chen and Manley(2003);Frontini et al. (2005)
TAF10	Taf25	TAF30	细胞周期;发育;组蛋白折叠	Metzger et al. (1999);Georgieva et al. 2000);Mohan et al. (2003)
TAF11	Taf40	TAF28	PIC 组装;激活因子特异性;组蛋白折叠	May et al. (1996);Caron et al. (1997);Kraemer et al. (2001)
TAF12	Taf61/68	TAF20/15	H2B-样;形成八聚体样结构;细胞周期;激活因子特异性	Reese and Green(2001);Selleck et al. (2001);Hamard et al. (2005)
TAF13	Taf19	TAF18	激活因子特异性;与 TAF11 相互作用;组蛋白折叠	Mengus et al. (1995);Lemaire and Collart(2000)
TAF14	Taf30		细胞周期;不重要的	Dahan and Kupiec(2004)
TAF15		TAF68	RNA 和 ssDNA 结合	Bertolotti et al. (1996)

表 12.4 Med 复合物亚基及其选取的功能

模块	因子	酵母	哺乳动物	功能	参考文献
头部	MED6	Med6	DRIP33、mMED6	发育;通用转录	Lee and Kim(1998);Gim et al. (2001)
	MED8	Med8	MED8	泛素化	Brower et al. (2002)
	MED11	Med11	mHSPC296	激活因子特异性激活	Han et al. (1999)
	MED17	Srb4	TRAP80、DRIP77	激活因子特异性;细胞活力	Koh et al. (1998);Boube et al. (2000);Lao et al. (2003)
	MED18	Srb5	p28b		
	MED19	Rox3	mLCMR1	应激应答基因表达	Evangelists et al. (1996);Singh et al. (2006)
	MED20	Srb2	TRFP		
	MED22	Srb6	Surf5	通用转录	Lee et al. (1998)
	MED28		Med23、Fksg20		
	MED29		Intersex	发育;抑制	Garrett-Engele et al. (2002);Wang et al. (2004)
	MED30		mTRAP25	通用转录	Baek et al. (2002)

续表

模块	因子	酵母	哺乳动物	功能	参考文献
中部	MED1	Med1	TRAP220、DRIP205、ARC205	与核受体相互作用；通用转录	Ito et al. (2000)；Ito and Roeder (2001)；Belakavadi and Fondell(2006)
	MED4	Med4	TRAP36、DRIP36，CRSP34	发育；通用转录	Balciunas et al. (2003)；Durr et al. (2006)
	MED5	Nut1		抑制；HAT 活性	Tabtiang and Herskowitz(1998)；Lorch et al. (2000)；Beve et al. (2005)
	MED7	Med7	MED7、DRIP34、CRSP33	发育；结构性铰链	Kwon et al. (1999)；Baumli et al. (2005)
	MED9	Cse2	Med25	激活因子特异性；抑制	Han et al. (1999，2001)；Kang et al. (2001)
	MED10	Med10	mNUT2	发育；抑制；激活因子特异性	Tabtiang and Herskowitz(1998)；Han et al. (1999)；Kwon et al. (1999)
	MED21	Srb7	TRAP19	抑制；发育；结构性铰链	Sun et al. (1998)；Tudor et al. (1999)；Gromoller and Lehming (2000)；Baumli et al. (2005)
尾部	MED2	Med2		通用转录；激活因子特异性	Hallberg et al. (2004)；Zhang et al. (2004)；van de Peppel et al. (2005)
	MED3	Med3		抑制	Papamichos-Chronakis et al. (2000)
	MED14	Rgr1	TRAP170、DRIP150、ARC150	激活因子特异性；抑制；通用转录	Sun et al. (1998)；Shim et al. (2002a)；Lau et al. (2003)；Toth et al. (2004)；Chen et al. (2006)
	MED15	Gal11	ARC150	激活因子特异性；通用转录	Han et al. (1999)；Park et al. (2000)；Sakurai and Fukasawa(2000)；Lu et al. (2002)；Fishburn et al. (2005)
	MED16	Sin4	TRAP95、DRIP92	通用转录；应激应答；激活因子特异性	Myers et al. (1999)；Reeves and Hahn(2003)；Kim et al. (2004)
CDK	CDK8	Srb10	SRB10、CDK8	抑制；Pol II CTD 磷酸化；PIC 支架	Hengartner et al. (1998)；Sun et al. (1998)；Akoulitchev et al. (2000)；Liu et al. (2004)；Pavri et al. (2005)；van de Peppel et al. (2005)；Andrau et al. (2006)；Zhu et al. (2006)
	CyclinC	Srb11	SRB11、Cyclin C	抑制	Sun et al. (1998)；van de Peppel et al. (2005)
	MED12	Srb8	TRAP230、DRIP240、ARC240	细胞分裂；激活因子特异性；发育	Treisman(2001)；Yoda et al. (2005)；Kim et al. (2006)；Rau et al. (2006)
	MED12L		TRALPUSH	类 MED12	

续表

模块	因子	酵母	哺乳动物	功能	参考文献
	MED13	Srb9	TRAP240、DRIP250、ARC250	细胞分裂;发育	Treisman(2001);Janody et al. (2003);Yoda et al. (2005)
	MED13L		THBAP2	类 MED13;发育	Muncke et al. (2003);Musante et al. (2004);Sato et al. (2004)
未知	MED23		TRAP150、DRIP130、Sur2	激活因子特异性;应激应答	Asada et al. (2002);Stevens et al. (2002);Lau et al. (2003);Kim et al. (2004);Mo et al. (2004);Wang et al. (2005)
	MED24		TRAP100、DRIP100	发育;通用转录	Ito et al. (2002);Dürr et al. (2006)
	MED25		TRAP93、ARC92	激活因子特异性	Mittler et al. (2003);Yang et al. (2004)
	MED26		DRIP70、CRSP70	活性 Med	Taatjes et al. (2002)
	MED27		TRAP37	发育	Dbrr et al. (2006)
	MED31	Soh1	mSoh1	通用转录	van de Peppel et al. (2005)

动力学研究

研究基本转录复合物的最初方法和说明复合物确实正在组装的最初生化论据，是基于粗制无细胞系统中的动力学研究（Hawley and Roeder 1985，1987）。当 HeLa 提取物和腺病毒晚期启动子（AdMLP）及核苷酸一起孵育时，在转录达到合成的峰值速率前有一个滞后期。相比之下，如果在无核苷酸存在的条件下，将模板和提取物预孵育大约 1 h，在加入核苷酸的瞬间，转录即达到峰值速率。这些研究暗示在转录中有一个的限速步骤，它可能归因于 PIC 组装所花的时间。预孵育过程中组装的复合物最初被称为“快速起始复合物”。像这些读出是 RNA 转录的简单时程研究，可在有激活因子和无激活因子的条件进行，并且对于那些想理解启动子转录机制的研究者来说，它们是一个很好的出发点（图 12.4）(如见 Johnson and Krasnow 1992)。当加上其他方法，如十二烷基肌氨酸钠考验（sarkosyl challenge）时，这些研究能产生激活因子功能相关的重要信息。的确，一旦转录分析方法到位，这些类型的实验都是相对简单的。因此，这类实验可以很快完成，并且产生有用的信息。

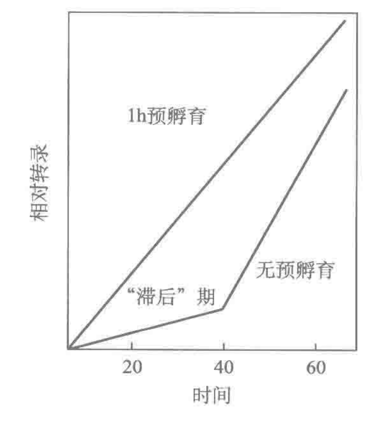


图 12.4 转录动力学。

十二烷基肌氨酸钠考验

十二烷基肌氨酸钠方法基于文献中的许多报道，这些报道表明在不同的组装或功能阶段中，原核转录复合物对破坏蛋白质-蛋白质和蛋白质-DNA 相互作用的试剂（如肝

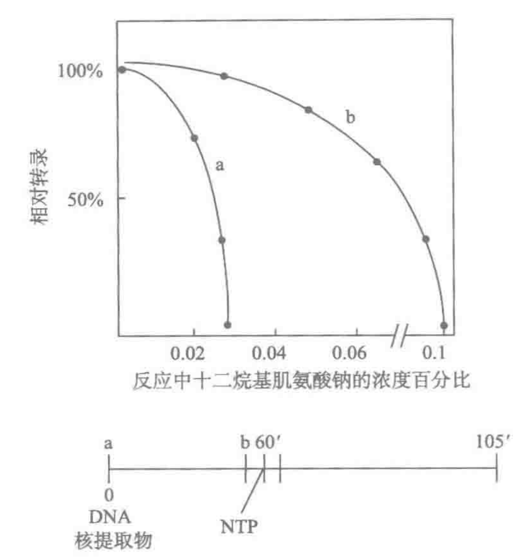


图 12.5　十二烷基肌氨酸钠抑制。

素）表现出不同的敏感性。Hawley 和 Roeder（1985，1987）发现离子型去污剂十二烷基肌氨酸钠可用于真核无细胞系统，以划定 Pol II 转录中的不同步骤，即起始、延伸和再起始之间的差异。简单地说，将模板和核提取物进行预孵育，然后在预孵育开始或结束时，或在加入核苷酸之后，加入各种浓度的十二烷基肌氨酸钠。当时的想法是转录的较早期步骤对较低浓度的去污剂敏感，因为当其他组分加入并且转录开始时，复合物变得更稳定（图 12.5）。如果和模板一起加入，低浓度的十二烷基肌氨酸钠（0.015%）就能够阻止转录（时间 a，图 12.5）。相比之下，在加入十二烷基肌氨酸钠之前，如果将模板加入到提取物中，并给时间使其组装成复合物，则转录对低浓度的去污剂具有抗性。但是，如果加入更高浓度的十二烷基肌氨酸钠（0.1%），即使在复合物组装后，它们也会抑制起始（时间 b，图 12.5）。最后，如果让聚合酶在有核苷酸的情况下起始并开始延伸，则转录甚至对 0.1%十二烷基肌氨酸钠也具有抗性。这些实验表明，十二烷基肌氨酸钠可用于区分转录过程中的不同步骤。

十二烷基肌氨酸钠考验研究已广泛用于区分转录中的不同阶段，并已应用到粗制系统（Zenzie-Gregory et al. 1992）和重建系统中，以区分含有不同子集纯因子的功能重要复合物（Lieberman 1994；Kim and Maniatis 1997）。其他类似的方法已用于探测基础转录复合物的差异灵敏度，包括增加的温度和盐浓度的利用（Cai and Luse 1987）。应当指出的是，许多早期的研究是在强基础启动子上进行。自然发生的细胞启动子还没有得到广泛检查。相同的步骤可以对十二烷基肌氨酸钠更敏感或更具有抗性，这取决于 PIC 对各启动子接触的强度。当使用相同浓度的十二烷基肌氨酸钠比较不同的启动子时，应考虑到这一点。

在粗制系统中，十二烷基肌氨酸钠考验通常用于将转录限制到单轮（即允许起始，但阻止再起始）。的确，单轮十二烷基肌氨酸钠考验实验可用于研究激活因子是否刺激转录复合物的组装。如果在有激活因子的单轮分析中观察到更高水平的转录，就可以暂时断定激活因子正在刺激复合物组装。

转录复合物组装的 DNase I 足迹和 EMSA 研究

有关基础转录复合物结合和构象的物理信息最初是通过核酸酶和化学足迹技术得到的。DNase I 和甲锭丙基 EDTA 足迹实验表明，高度分级分离的 TFIID 与 TATA 盒相结合，并成为其他通用转录因子结合的核心。这些研究也表明激活因子能影响 TFIID 向 TATA 盒的结合，以及 TFIID 在 TATA 盒上的构象，并为激活因子募集含 TFIID 复合物的未来研究提供了架构（Sawadogo and Roeder 1985a；Horikoshi et al. 1988；

Nakajima et al. 1988；Van Dyke et al. 1988）。

TFIID 的 TBP 亚基的克隆和其他通用转录因子的进一步纯化，让基础复合物的更精细分析成为可能。TBP 纯蛋白的小尺寸和可得性使得利用 EMSA 和 DNase I 足迹研究基础转录复合物成为可能。利用这些技术，建立了基础转录复合物组装的逐步路径（综述见 Orphanides and Reinberg 2002）。这个贡献为后来活化转录的研究提供了重要且关键的框架。

最初的实验中使用了含有 AdMLP 的^{32}P 标记限制性片段（Buratowski et al. 1989），放射性标记片段与单独的 TBP 或分级分离的通用因子的不同组合一起孵育，复合物用非变性聚丙烯酰胺凝胶电泳（即 EMSA）分离。为了说明该方法，考虑一下 TBP 和 TATA 盒的结合：单独的 TBP 与 TATA 盒结合很差，并且产生快速迁移的复合物。加入 TFIIA（自身不与 DNA 结合）时，强烈刺激 TBP 结合，形成所谓的 TA 亚复合物。TA-DNA 复合物比 TBP-DNA 复合物迁移得更慢，说明 TFIIA 确实加入了复合物。通过改变各因子的组合并检测所形成的复合物，就可以确凿地建立逐步组装路径。

通过将 EMSA 与 DNase I 足迹结合（图 12.6）可以推断出启动子中各因子结合位置的低分辨率信息。一旦复合物形成，就可以在凝胶电泳之前利用 DNase I 对其进行轻度消化。用 EDTA 抑制进行进一步的切割，则留下仅在^{32}P 标记探针的一条单链上带有切刻的 DNA，而另一条链保持完整。然后，复合物通过 EMSA 分离。因为 DNA 沿单链被切刻，在自然凝胶上它不会与蛋白质解离（Buratowski et al. 1989）。放射自显影后，从凝胶上切下被切刻的 DNA，纯化除去蛋白质，然后在聚丙烯酰胺/尿素测序胶上分析，显示出各个复合物的 DNase I 足迹。该方法对于活化转录有很多应用，应牢记它，因为稍后我们要讨论这一问题。

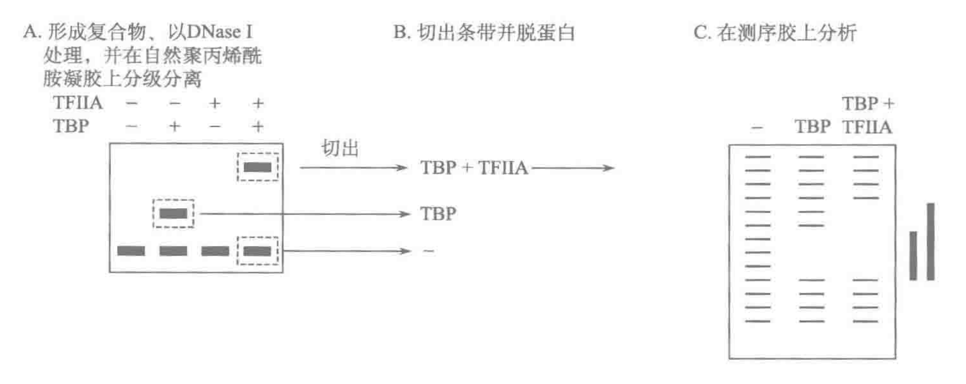


图 12.6 EMSA 复合物及足迹示意图。

尽管 Buratowski 和 Sharp 的早期研究确立了方法学及最初的实验步骤，但使用的因子制剂有些粗糙，不能用于明确地划分组装路径中的晚期步骤。后来的研究使用了纯化的或重组的通用因子，确立了最终的通路（图 12.7）(Orphanides and Reinberg 2002；Hahn 2004)。在第一个步骤中，TBP 与 TATA 结合，此反应受 TFIIA（或者在有些实验中是 TFIIB）的辅助和强化。TA 亚复合物招募 TFIIB（TAB 复合物），TFIIB 顺次招募与 TFI-

IF（TABpolF）稳定结合的未磷酸化 IIA 型 RNA Pol II。一旦聚合酶结合，它就募集 TFIIE，之后 TFIIE 招募 TFIIH，形成 TABpolEFH 复合物。在 ATP 存在的情况下，TFIIH 磷酸化 Pol IIA，产生 Pol IIO 型 RNA Pol II。ATP 的存在也诱导 DNA 在一个需要 TFIIH 的反应中发生熔解（如下文所述），需要提醒的是，虽然上述组装步骤示意图从热动力学角度来说是最可能的路径，但在与 Reinberg 实验室所使用的不同的实验条件下，其他途径的存在也是可能的。

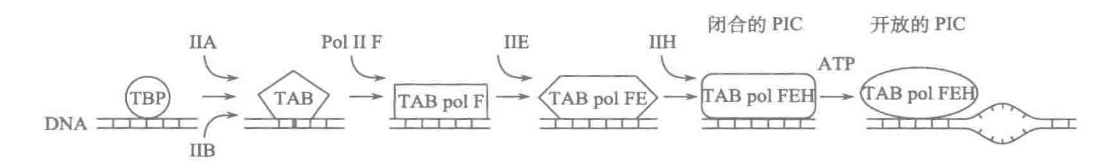


图 12.7 转录复合物组装步骤示意图。

光交联

虽然 DNase I 足迹法可取得因子在 DNA 上定位的信息，但足迹的分辨率有些低，对什么蛋白质与 DNA 片段的哪些部位相接触并不总是十分清楚。例如，当复合物正从游离的因子进行组装时，随着因子的加入，会有新的足迹出现。新加入的因子的直接结合并不一定造成新的足迹，相反，新的因子会诱导已结合的因子发生构象变化或重排。因此，使用具有光活化性的交联剂是必不可少的，以提供在基础转录复合物中发生的直接的蛋白质-DNA 相互作用图景，特别是探测复合物中的构象变化。激活因子怎样诱导这些变化至今鲜为人知，很多共激活因子（包括 Med）还没有广泛地在本分析方法中被测试。因此，它仍然是一个适用于未来激活因子-共激活因子相互作用研究的有趣的技术。

两种通用的基于 DNA 的交联方法已经用于 Pol II 复合物。第一种方法使用的是 5-[*N*-（*p*-叠氮苯甲酰基）-3-氨基烯丙基] -dUTP（N_3RdUTP）（图 12.8A）。N_3RdUTP 交联首先用于 RNA Pol III 系统中的转录复合物组装（Bartholomew et al. 1991, 1993）。N_3RdUTP 在距离掺入 DNA 替代胸腺嘧啶的尿嘧啶环 9～10Å 处安置一个具有光反应特性的叠氮基团，该试剂探测与 DNA 大沟发生相互作用的蛋白质。

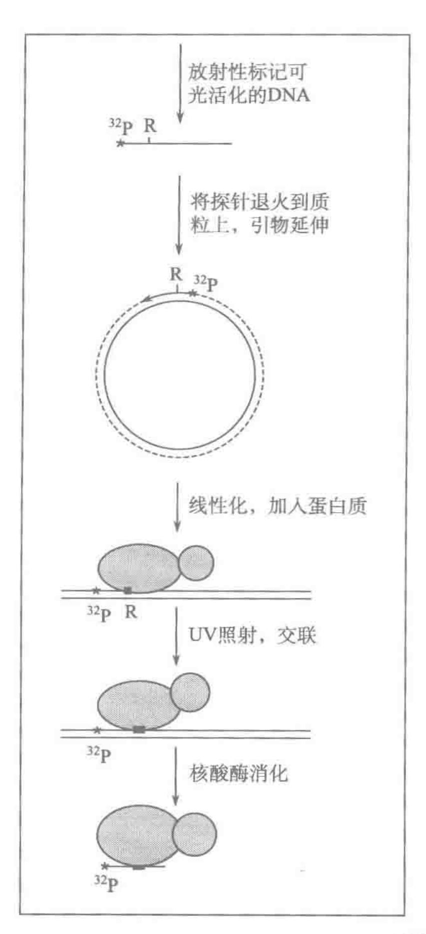
第二种方法涉及在硫代磷酸衍生的磷酸盐上连接一个苯基叠氮基团（图 12.8B）（Lagrange et al . 1996，Kim et al. 1997）。该方法在距离磷酸骨架 9.7Å 处安置一个能被光活化的基团，原则上用于探测发生在大沟和小沟中的相互作用。交联基团被加在 ^{32}P末端标记的寡核苷酸上。寡核苷酸和单链模板杂交，并利用 DNA 聚合酶进行引物延伸，以产生双链 DNA 分子。通过限制性内切核酸酶消化和凝胶分离去除标记的衍生启动子片段。

一旦准备好衍生片段，不同组合的重组因子就与之结合，产生的复合物经受 UV 照射处理（312～350 nm）。UV 处理活化叠氮基团，产生氮烯，这是一个能将 DNA 交联到蛋白质上的具有高度反应性的基团。之后，蛋白质-DNA 复合物经受 DNase I 广泛消化。DNase I 切割被交联 DNA 的距蛋白质几个核苷酸远的所有核苷酸。^{32}P 标记一般位于靠近交联基团的部位，因此，DNase I 处理后，仍有与蛋白质结合的少量 DNA 含有

A

N_3R dUTP

B



叠氮基苯

9.7□

图 12.8 蛋白质与 DNA 的交联。(A) RNA Pol III 系统中转录复合物组装的 N_3RdUTP 交联（经许可改编自 Bartholomew et al. 1991，© Society for Microbiology)。(B) 硫代磷酸衍生的磷酸盐上连接一个苯基叠氮基团的交联（经许可改编自 Bartholomew et al. 1991；Lagrange et al. 1996。© National Academy of Sciences，U. S. A.）

^{32}P标记。放射性标记使蛋白质在 SDS-聚丙烯酸胺凝胶上自显影检测。因为通用因子是重组的或高纯度的，其亚基的大小已知，所以可用于确定被标记蛋白质的身份。

数项研究说明了这些技术在评价基础复合物组装中的应用。在其中一项研究中，光交联法证明了 Pol II 最大的两个亚基形成一个延伸 220Å 的通道，该通道与 AdMLP DNA 起始位点上、下游相互作用（Lagrange et al. 1996）。该项技术也用于研究 TFIID 的结合及 TAF_{II}沿 AdMLP 的定位，结果表明，TFIIA 在 TFIID 与 DNA 的结合中诱导关键性的构象变化，这是一种效应，它暗示 TFIIA 在活化转录中的机制（Oeigeschlager et al. 1996）。PIC 与 AdMLP DNA 的高分辨率的光交联被用于模仿复杂的蛋白质-DNA 相互作用（图 12.8）（Forget et al. 2004）。对于溶液中进行的光交联研究有一个警告，那就是可能存在几个亚复合物。为了解决这个问题，研究人员在溶液中组装了一个由 TBP、TFIIB、TFIIF、TFIIE 和 Pol II 构成的复合物，接着进行聚丙烯酰胺凝胶电泳分离和紫外线光交联。从凝胶中切下仅代表完整 PIC 的条带，并且通过 SDS-PAGE 和放射自显影观察光交联的多肽（图 12.9）。交联数据支持一种模型，据此启动子 DNA 包裹在 Pol II 夹钳（clamp）周围，从而让 TFIIE 和 TFIIF（与夹钳缔合）接触上游和下游元件。正如这项研究所表明的，这些光交联研究的重要方面是，它们可与各种 PIC 亚复合物的结构发生相互关系，就像它们在文献中所显示的那样。

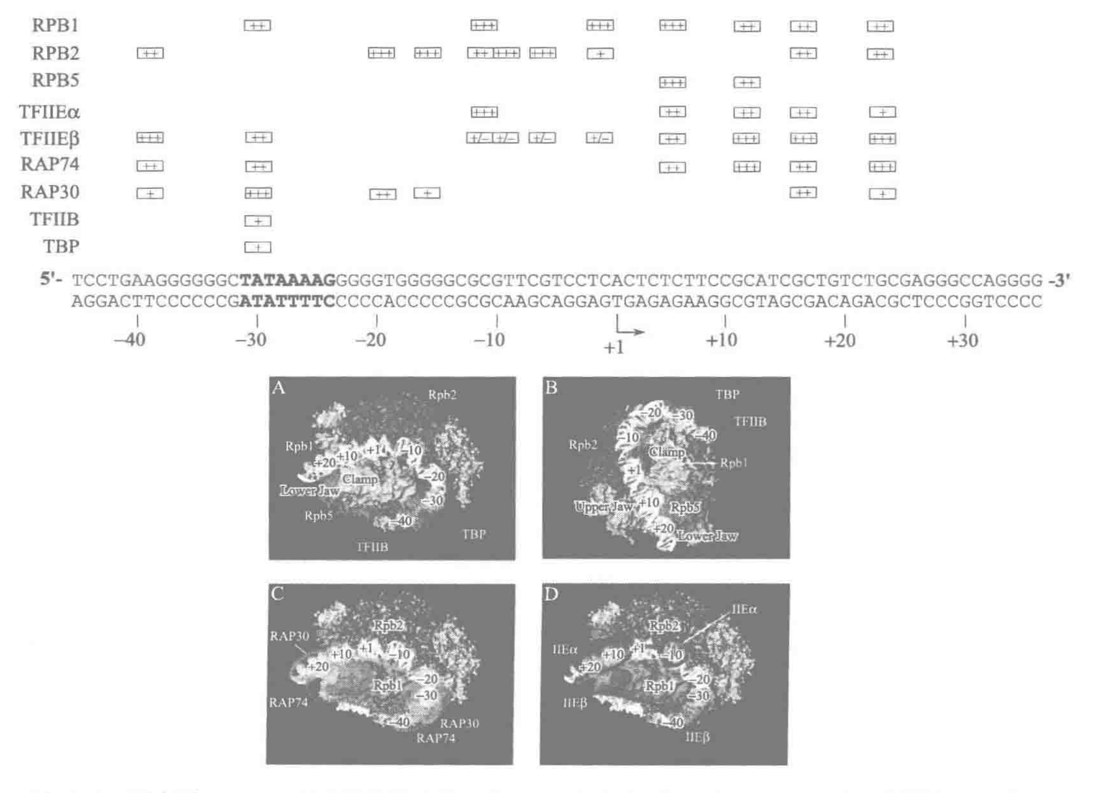


图 12.9　蛋白质-DNA 在启动子处的交联。交联总结（上）和一个基于 Pol II 和 TFIIB-TBP 结构及交联数据的模型（下）（经许可转载自 Forget et al. 2004，© American Society for Microbiology）。

光交联也可用于研究转录复合物内蛋白质-蛋白质的交联。该技术用于研究大肠杆菌中代谢物基因激活因子蛋白与 RNA 聚合酶亚基之间的接触（Y. Chen et al. 1994），随后又应用于真核系统。Hahn 实验室采用此策略鉴定酵母提取物中 GCN4 和 GAL4 激活结构域的靶标。在这些研究中，半胱氨酸被放置在两个激活因子的激活结构域的不重要的部分。不重要的残基是早期使用一些策略通过突变鉴定的，如第 10 章中就叙述了这类策略。光交联剂 *N*-（（2-吡啶硫）乙基）-4-叠氮水杨酰胺（PEAS）用^{125}I 标记。PEAS 是一个芳基叠氮，它经历了自己的吡啶二硫化物基团与蛋白质半胱氨酸上巯基基团间的二硫化物-巯基交换。一旦紫外线照射，叠氮将与 14Å 距离内的邻近蛋白质交联，形成共价键连接（图 12.10）。通过交联法与固定化模板分析（如下所述）的结合，研究者发现，GAL4 和 GCN4 接触组装自酵母提取物的 PIC 的非常特异的亚基。例如，GCN4 与 TRA1、GAL11 与 TAF12 交联（Fishburn et al. 2005），然后，它们就能利用各种酵母突变体的提取物来评价相互作用的相关性。该技术还可以用于研究组成 PIC 的通用转录因子间的相互作用。PIC 上 TFIIB 的光交联确认了 TFIIB 与 Pol II 亚基 Rpb1 和 Rpb2 的相互作用，同时揭示了 TFIIB 与 TFIIF 亚基 Tfg1 和 Tfg2 的相互作用，这表明 TFIIB 在稳定 Pol II-TFIIF 相互作用中发挥作用（Chen and Hahn 2004）。

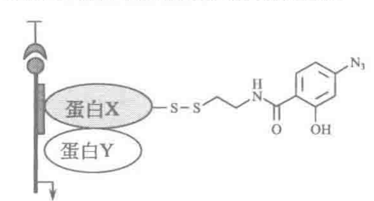
N-((2-吡啶硫)乙基)-4-叠氮水杨酰胺(PEAS)

1.放射性标记PEAS

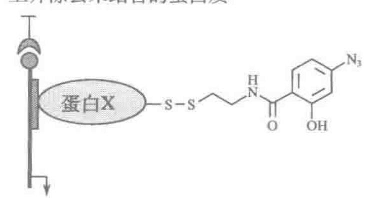
2. 将PEAS与蛋白质孵育

蛋白X

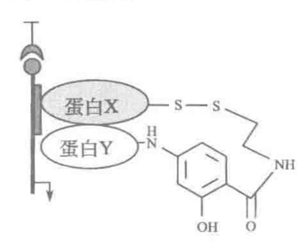
4. 加入核提取物并除去未结合的蛋白质



3. 将蛋白质预结合到固定化DNA模板上并除去未结合的蛋白质



5. 以UV光活化



6. 加入DTT并用SDS-PAGE分析

图 12.10　PEAS 光活化交联剂介导的蛋白质-蛋白质交联（改编自 Cai et al. 2001，© National Academy of Sciences，U. S. A.）

通用机器的结构-功能分析

通用转录机器的诱变和结晶学分析为真核转录设备提供了其他的机制性细节。最全

面的方法将上述的各种分析方法与诱变相结合，来分析参与相互作用的各个因子的表面。迄今为止，Pol II、TFIIA、TFIIB、TBP和几个TAF都已被结晶，其结构也已被解析。TFIIF和TFIIH单个亚基的结构已经被解析。很多其他因子（如TFIID和Med）的冷冻电镜结构在较低分辨率水平上也已经被解析。在许多情况下，诱变分析结果与晶体结构分析结果完美吻合，而有些情况下则出现差异。最清楚的理解来自各种分析的汇集，这些分析包括结构分析、诱变分析和重组转录系统中突变体的功能分析，其中最重要的是重组转录系统中突变体的功能分析，而这个问题涵盖在第10章中。

开放复合物的形成、起始和启动子逃脱

ATP类似物和一个能量依赖性步骤

在原核细胞和真核细胞中，转录都遵循概念上相似的途径。聚合酶及其相关因子在启动子DNA上形成一个关闭复合物，随后形成一个开放复合物，在开放复合物中包含起始位点的DNA被解链。不像常见的细菌σ^{70}系统，Pol II（不是Pol I或Pol III）机器需要ATP才能形成开放复合物，然后聚合酶起始转录，并最终从启动子上逃脱并延伸。有很好的证据表明，Pol II就像其细菌对应物，会遭受流产性起始的反复循环，在这种流产性起始中，聚合酶起始转录，但不能清除启动子。

体外转录系统发展于20世纪80年代初，这种系统的最早应用之一是鉴定ATP水解的重要能学作用（Bunick et al. 1982）。在该研究中，在提取物中孵育并分析由Pol II转录的AdMLP和E4启动子及由Pol III转录的VA基因，以确定在不能进行水解的ATP类似物AMP-PNP存在的情况下，转录是否会发生。AMP-PNP在β磷原子和γ磷原子之间含有一个亚氨基基团。众所周知，转录需要水解核苷酸的α-β键，含有完整α-β键的AMP-PNP能够支持由细菌RNA聚合酶负责的转录起始和延伸。尽管Pol III负责的转录不受AMP-PNP的影响，但Pol II负责的转录被完全废止。其他β-γ酰亚胺衍生物（包括GMP-PNP）对转录没有影响，并且在ATP存在时能够完全支持Pol II的延伸。如果用dATP代替ATP加入混合物中，在AMP-PNP存在时Pol II转录发生。由此得出结论，即dATP能提供一个可水解的β-γ键，它满足能量的需求。该能量使Pol II起始转录，并通过利用AMP-PNP的完整α-β键进行延伸。这一重要发现开启了近15年的阐明Pol II转录中ATP能学作用的研究。

随后的动力学研究检查了ATP β-γ键水解的作用（Sawadogo and Roeder 1984; Luse and Jacob 1987）。总的来说，这些数据指向Pol II开放复合物是ATP β-γ键需求的靶标。这一观点更直接的支持最终来自一项研究，该研究表明超螺旋而非线形AdMLP能够避开对ATP β-γ键水解的需求和对重组转录系统中通用因子TFIIH的需求（Timmers 1994）。众所周知，负超螺旋促进DNA解链，该事实表明，开放复合物的形成确实是ATP水解影响的步骤。此外，观察到TFIIH被绕开就意味着正是该因子实现了解链的步骤。众所周知，TFIIH含有DNA解旋酶和激酶亚基，两者都利用ATP（综述见Orphanides et al. 1996; Hahn 2004），因此推测这些活性之一负责DNA解链。

一篇PIC上TFIIH活性的综述支持TFIIH导致DNA解链的假说（Dvir et al. 2001）。两个独立的实验室利用AdMLP上的DNA交联，观察了在GTF及pol II存在的情况下

TFIIH 在开放复合物上的结合（Douziech et al. 2000；Kim et al. 2000）。在这些系统中，观察到开放复合物的形成是 ATP 和 TFIIH 依赖性的。同时观察到 TFIIH 与上游（−29～−45 bp）、近转录起始位点（−5 bp）和下游（+13～+38 bp）DNA 的交联（Douziech et al. 2000）。TFIIH 的着色性干皮病（DNA 修复酶缺乏病）B（XPB）解旋酶亚基中的突变消除了 DNA 熔解，从而使该亚基牵扯进这个过程。交联和诱变数据导致提出这样的建议，即 TFIIH 结合引起其结合部位 DNA 的较大的负超螺旋缠绕，从而导致转录起始位点周围 DNA 的轻微解旋。这一解旋对 XPB 解旋酶来说很显著，足以使其之后进一步以依赖 ATP 的方式解旋 DNA。Kim 等（2000）发现 TFIIH 只交联到下游的 DNA（+3～+25 bp）上。这导致他们提出一个不同的模式，即上游 DNA 是由蛋白质相互作用固定，而 TFIIH 造成转录起始位点下游 DNA 的旋转。这种下游 DNA 的旋转（或扭转）加上上游 DNA 的固定，促使转录泡的形成。

高锰酸盐探测

高锰酸盐敏感性是一个直接研究开放复合物形成的很好的方法。该方法最初用于测定 *E. coli* RNA 聚合酶形成的开放复合物（SasseDwight and Gralla 1989），随后被用于检测 RNA Pol III（Kassavetis et al. 1992）和 Pol II（Wang et al. 1992a）形成的开放复合物。高锰酸钾选择性地修饰单链，而非双链 DNA 中的胸腺嘧啶（T）（图 12.11）。当起始位点解链时，T 残基对高锰酸盐的修饰变得敏感，形成胸苷乙二醇。被修饰的 T 能阻止引物延伸反应中由 *Taq* 酶引起的延伸（见下文），而且对哌啶切割敏感。延伸终止产物或哌啶切割产物可在聚丙烯酰胺/尿素凝胶上进行电泳分离，并作为开放复合物形成的指示物。Gralla 实验室利用 HeLa 核提取物进行高锰酸盐敏

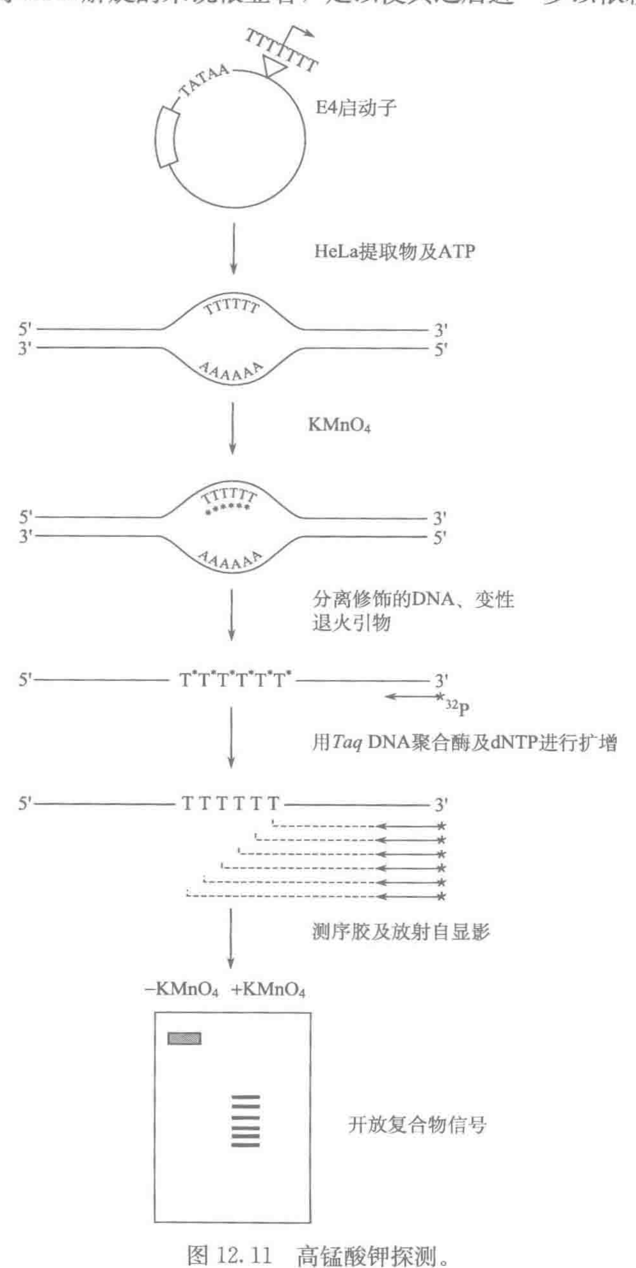


图 12.11 高锰酸钾探测。

感性检测，结果表明，腺病毒 E4 起始位点解链形成 Pol II 开放复合物需要 ATP β-γ 键。该技术（方案 12.7）在研究活化转录中也大有用途，对此后面有更详细的描述。

预熔解模板

Roeder、Luse、Grana 和其他研究人员的研究产生了一个假说，即开放复合物形成期间，ATP 水解驱动 DNA 解链。此外，Timmers 研究提示无 ATP 存在时，超螺旋的自由能能代替 β-γ 键水解。该假说的一个预测是，预熔解的 DNA 模板也应能绕过对 ATP 和解链起始位点的因子的需要。Tantin 和 Carey（1994）构建了一个在起始位点上含有 10 bp 错配的腺病毒 E4 模板，并且表明该模板确实绕过了活化转录期间对 ATP β-γ 磷酸二酯键的需要。相同的预熔解 E4 模板也能在含纯因子的基础转录系统中绕过对 TFIIE 和 TFIIH 的需要（Holstege et al. 1995），其他研究人员利用相似的方法也独立地得出了这一结论（Pan and Greenblatt 1994）。剩下的一个问题是：激活因子是否能调控解链步骤。在高纯度重建系统中进行的研究表明，预熔解模板上的活化转录仍然需要激活因子 GAL4-VP16。但是，在活化系统中，预熔解模板仅仅绕过了对 TFIIH 的需要，而不能绕过对 TFIIE 的需要（Tantin et al. 1996），这一结果提示活化转录所必需的其他蛋白质（TAF 或共激活因子）迫使活化转录需要 TFIIE。

Conaway 实验室进行的研究（Dvir et al. 1996，1997a，b；Moreland et al. 1999；Spangler et al. 2001）使用了预熔解模板，以确定有效的启动子逃脱依赖于 TFIIH、ATP 及下游 DNA（距 AdMLP 启动子 39～50 bp）。此外，他们还表明只有当 TFIIH 和 ATP 也是必不可少的时候，转录起始才依赖于下游 DNA 序列（距 AdMLP 启动子 23～39 bp）。最后，他们发现，下游 DNA 序列对于预熔解模板上的起始和（或）启动子逃脱是不必需的（即 TFIIH 不存在），这表明 TFIIH 需要下游的 DNA，以促进启动子解链（综述，见 Dvir et al. 2001）。

向延伸的过渡

启动子逃脱，即从开放复合物向生产性延伸的过渡是基因转录中的一个关键步骤。起始后，聚合酶有下列三种选择，包括：①它可以延伸短距离（7 个核苷酸左右），但不从基础转录机器上释放（流产性起始）；②也可以进一步延伸（距离起始位点 100 bp 以内）并暂停（启动子近端暂停；见第 3 章）；③它还可以继续生产性延伸。研究提示所有这些步骤都受激活因子的调控（综述见 Bentley 1995；Svejstrup 2004）。的确，激活因子（如 GAL4-VP16）在模式系统中既刺激启动子逃脱，也阻止启动子近端暂停（Kumar et al. 1998）。这不可能是模式系统中的人工产物，因为已发现有些天然激活因子（如 HIV Tat 蛋白）在体内也阻止启动子近端暂停（综述见 Yankulov and Bentley 1998；Sims et al. 2004a）。

当前，虽然 TFIIH 和 TAK 复合物已经涉及向延伸过渡的过程，但这个特殊步骤的相关细节仍不清楚（综述见 Jones 1997）（见第 3 章）。在转录起始向延伸的过渡中一个关键的事件是 Pol II 羧基端结构域（CTD）尾巴的磷酸化，该磷酸化反应可以被 P-TEFb 调节（Price 2000）。在秀丽隐杆线虫（*Caenorhabditis elegans*）中，P-TEFb 亚基 Cdk9 对 Pol II CTD 丝氨酸-2 的磷酸化是必要的（Shim et al. 2002b）。此外，使用缺失的 HeLa 提取物的体外转录表明，P-TEFb/Brd4 相互作用对获得转录的最大激活是必不可少的（Yang et al. 2005）。重要的是要注意 P-TEFb 也有其他靶标，包括某些延

伸因子，并且，这些因子的磷酸化比CTD磷酸化可能发挥更重要的作用。对从启动子近端暂停到延伸的过渡的详尽考察超出了本章的范围。相反，我们讨论TFIIH在启动子清除中的作用，并用一些挑选的研究来说明这一问题是如何被研究的。重点是激活因子确实介入了从起始到启动子逃脱的过渡。生化分析能使研究人员探测该机制。

固定化模板法最初被发展的目的是用于纯化Pol III转录复合物（Kasher et al. 1986），随后再应用于分析Pol II复合物（Arias and Dynan 1989）。在一项精美的研究转录复合物开始生产性延伸时其组分的实验中（Zawel et al. 1995），将生物素标记的启动子片段固定在链亲和素珠子上，并加入通用因子（图12.12）。漂洗除去未结合的

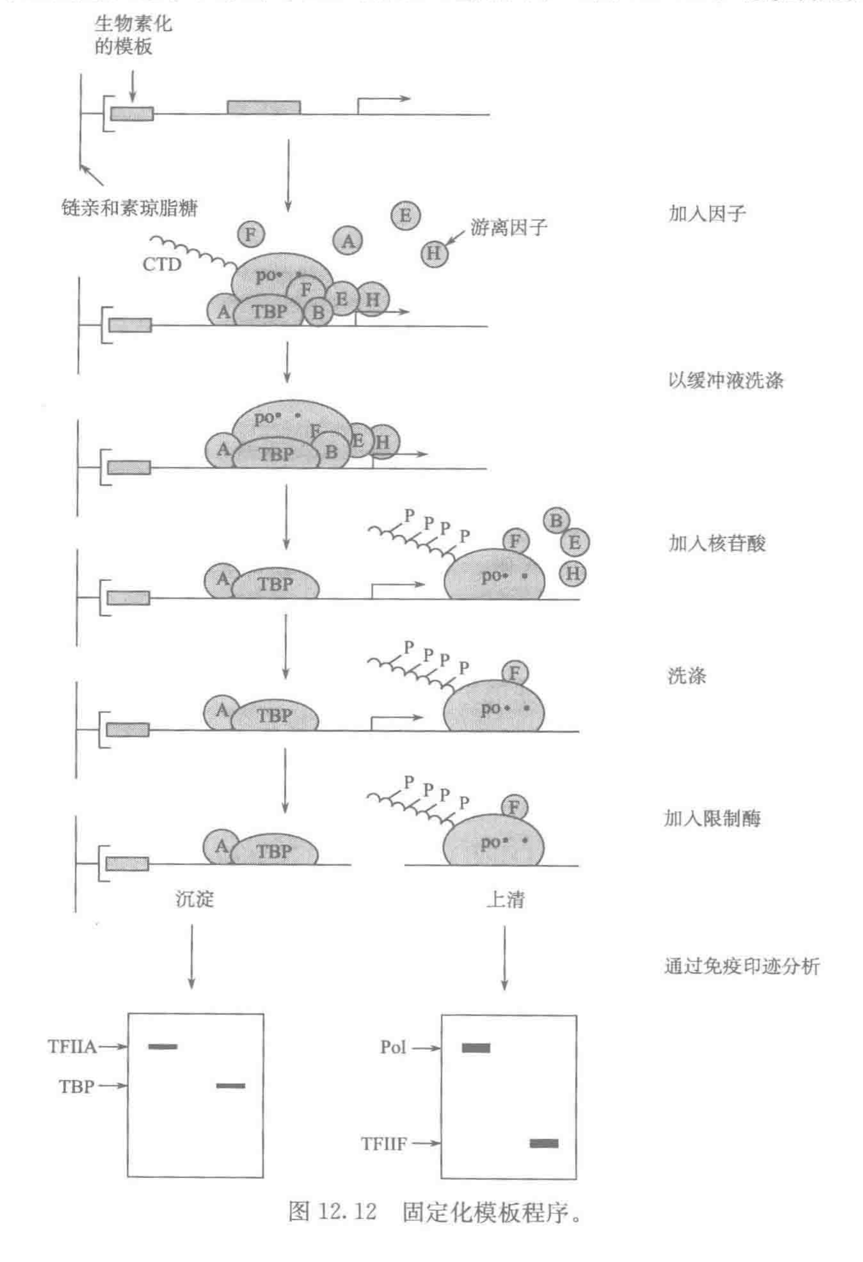


图12.12　固定化模板程序。

因子后，加入不同组合的核苷酸，以控制延伸的程度。产生的延伸复合物再一次漂洗，以去除延伸过程中解离下来的因子。核心启动子间的限制性位点使研究人员能够释放延伸复合物，并通过免疫印迹检查它们的组分，将该组分与其在启动子上结合的复合物的组分进行比较，结果发现，起始和延伸后，TFIID（即 TBP）被保留在后面，而 TFI-IB、TFIIE、TFIIF 和 TFIIH 被释放。TFIIE 在第 10 个磷酸二酯键形成之前释放，而 TFIIH 在转录到＋30 后释放。延伸期间，TFIIF 能保持与 Pol II 至少为瞬时性的缔合，事实上 TFIIF 能够解离下来，并重新加入暂停的聚合酶。

已有多项重要研究考察了启动子逃脱的机制和影响启动子逃脱的因素。Goodrich 和 Tjian（1994）首先提示了 TFIIH 参与启动子逃脱，后来的研究证实并拓展了这一发现。这些研究使用了在起始位点处预熔解的 AdMLP 模板，因为这种模板能绕过起始对 TFIIH 的需要，起始后可以往反应体系中加入 TFIIH，以探测其对启动子逃脱的影响（Dvir et al. 1997a）。为了研究在预熔解模板上的起始，可以将与起始位点互补的二核苷酸作为 Pol II 的引物（Samuels et al. 1984）。Pol II 往引物上添加一个放射性标记的核苷酸，并且该反应的效率被视为是对起始的一种测量。虽然 TFIIH 对预熔解模板上二核苷酸的引发效率几乎没有影响，但当加入一整套核苷酸时，它能明显增加全长转录本的生产。结果表明，TFIIH 除了在起始位点解链中的作用外，Pol II 从启动子上的有效释放也需要 TFIIH。延伸反应对 CTD 激酶抑制剂敏感性的缺乏，说明启动子逃脱需要的是 TFIIH 的 DNA 解旋酶活性，而非其激酶活性。

Holstege 等（1997）进一步在结构水平上研究了延伸过渡。用高锰酸盐的敏感性探测当 Pol II 开始延伸时开放复合物中的结构性变化。用 3′-*O*-Me-GTP 控制延伸过渡，3′-*O*-Me-GTP 是一个链终止 GTP 类似物，当模板上在起始位点下游的各个位置含有 G 残基时，3′-*O*-Me-GTP 就会掺入生长的 mRNA 链，从而导致 Pol II 停止作用，终止链延伸。稳定的生产性开放复合物的形成需要 ATP 水解和通用因子，包括 TFIIE 和 TFIIH。起始开放复合物（从－9 到＋1）可被 ATPγS 逆转，ATPγS 是一个 ATP 类似物，可与 ATP 竞争，但无水解能力。当加入 NTP 时，转录泡扩大，直到 9 核苷酸 mRNA 产生，在这一点上，转录泡的上游边缘消失，从而提示在起始阶段该复合物经历了一个关键的过渡。最近的工作表明，虽然有令人信服的数据表明 DNA 泡的大小和上游泡的消失提供了对 TFIIH 需求的终点和启动子逃脱的信号（Pal et al. 2005），但正是转录本的 8 核苷酸本身，而不是在启动子上的相对位置，决定了过渡（Hieb et al. 2006）。XPB 的晶体结构论证了 TFIIH 的 XPB 解旋酶如何控制 DNA 在启动子处熔解的特定模式（Fan et al. 2006）。

激活因子可差异调控启动子逃脱的思想来自几个生化实验。在 TFIIH 缺乏的情况下，形成二核苷酸引发的流产性短 RNA，并且 Pol II 的延伸在起始位点下游＋12～＋17处停止。加入 TFIIH 和 ATP 能刺激全长转录本的生产，但仍能观察到短 RNA。但是，加入激活因子 GAL4-VP16 将导致短转录本减少，并增加全长转录本的生产，说明激活因子确实能刺激启动子逃脱（Kumar et al. 1998），这一结果支持以前已报道的体内结果（如见 Bentley 1995）。在过去的几年中，已经表明很多激活因子与 TFIIH 直接相互作用，并刺激转录，包括 p53 和 VP16 及几个核受体。的确，已经出现了一些具有 TFI-IH p62 亚基的激活结构域的结构（Di Lello et al. 2006）。因此，有可能 TFIIH 是激活因子

真实的靶标，而且 DNA 熔解和（或）启动子逃脱将是合情合理的调控检测点。

活化复合物在启动子上的组装

到此，我们已经主要讨论了基本复合物及延伸的机制和结构，尽管我们已触及活化的各个方面。随后的几节涵盖了数个分析方法，已证明这些分析方法可为理解激活因子在 PIC 组装上的生化作用及功能提供信息。应当指出的是，这些分析方法中的很多方法可引出不同类型的信息。此外，因为这些分析方法涉及一些新的重要技能的获得，因此具有很好的理由去追求它们。在这些分析中，我们推荐固定化模板分析，因为它在技术上简单而强大，尽管掌握这一技术需要一些经验和实践。

固定化模板法

固定化模板分析是研究激活因子对 PIC 组装影响的功能最强大、信息最丰富的分析方法之一；它也可以用于研究抑制。的确，该方法的简单性使得它成为研究几乎任何启动子的一个很好的起点，并且在游离模板和染色质模板上它都可以使用（见方案 12.6）。该方法可分析具有转录活性的 PIC 的组分。在一个典型的分析中，启动子生物素化并连接到链亲和素偶联的磁珠上，激活因子与固定模板预结合，紧接着洗涤，然后添加 HeLa 粗提物或纯化的共激活因子（图 12.13）。

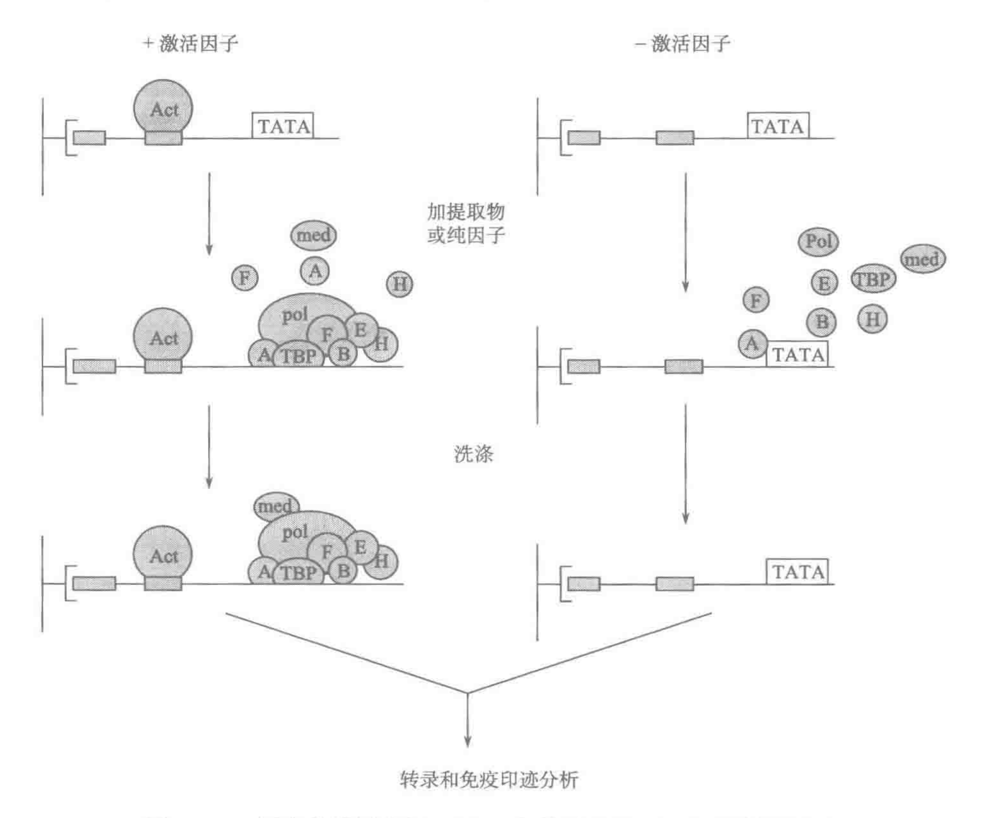


图 12.13　固定化模板分析：（+）有激活因子；（−）无激活因子。

在一项由 Black 等（2006）进行的研究中，GAL4-VP16 被组装到 GAL4 应答模板上（图 12.14A）。GAL4-VP16 结合后，加入 HeLa 提取物并孵育不同时间，结果发现各种因子的结合（图 12.14B）与模板的转录相关（结合研究使用代表各种复合物的抗体进行免疫印迹，因为对所有复合物的所有亚基都进行印迹是行不通的，原因在于某些抗体资源缺乏、成本较高并会造成冗余）。该研究表明，复合物具有转录活性（图 12.14C），并且免疫印迹揭示 GAL4-VP16 与 GTF 及 Pol II 一起招募两个主要的共激活因子——TFIID 和 Med（图 12.14B）。重要的是，动力学研究表明，PIC 组装的完成与转录的起始非常一致，可在几乎任何模板上用任何激活因子进行这种类型的分析。

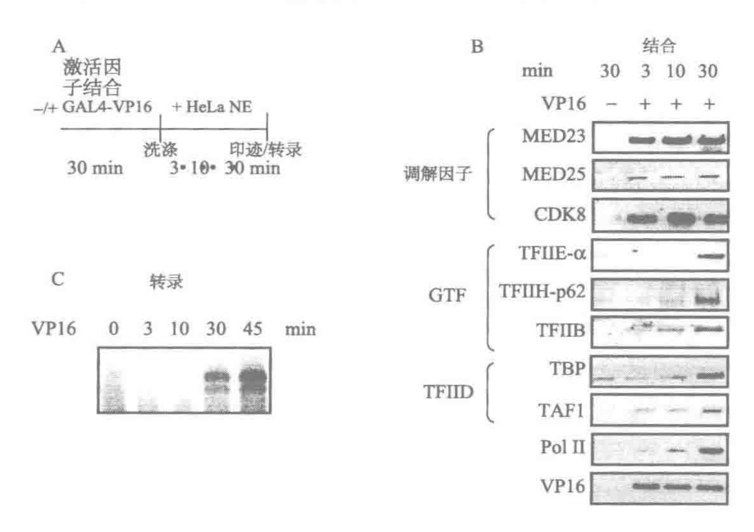


图 12.14 GAL4-VP16 指导的辅因子向固定化 DNA 和染色质的有序募集。（A）固定化模板分析时线示意图。30 min 激活因子结合期之后 HeLa 核提取物（NE）孵育 3 min、10 min 或 30 min。在缺乏 Gal4-VP16 的情况下，为 30 min 孵育做一个对照反应。（B）HeLa 核提取物（NE）孵育后，固定化 DNA 模板结合蛋白的免疫印迹。（C）PIC 组装完成后体外转录发生。DNA 模板和 HeLa 核提取物孵育方式与（B）中相同，但随核提取物加入 NTP 并且让转录进行到标示的时间（经许可转载自 Black et al. 2006，© Elsevier）。

该分析还可用于研究纯化的复合物。在那些提示 TFIID 和 Med 相互作用的资料的基础上，Johnson 和 Carey（2003）纯化了两个主要共激活因子复合物，并表明它们协作结合固定化启动子。当补充缺失 Med 和 TFIID 的 HeLa 提取物时，启动子 DNA 上 TFIID 和 Med 的预组装（DAMed 复合物）刺激转录的速度。他们断定 DAMed 复合物的组装是转录中的一个限速步骤。的确，使用相同类型的缺失 HeLa 提取物，GAL4-VP16 单独与 Med 或 TFIID 都不足以募集 GTF。相反，GAL4-VP16、TFIID 和 Med 的并合面（combined surface）导致有效的 PIC 组装。

为了研究转录复合物组装的机制和动力学，Ranish 等（1999）从野生型或突变型细胞中制备了酵母提取物，其中突变型细胞在 TBP、TFIIA、TFIIB、Pol II 或几个调解因子亚基中含有条件性突变，由这些提取物组装了复合物。通过分析野生型和突变型

复合物的转录活性测定了 GAL4-AH 对复合物组装的影响，并且通过免疫印迹分析了复合物的因子组分。结果表明，TFIID-TFIIA 复合物（见第 1 章）在一步中结合，而 Med 和 GTF 在随后的一步中结合。在 TFIIA 和 TFIID 结合后才结合的任何组分（如 TFIIB、Med 和 Pol II）中的突变，都影响所有后续因子（如 TFIIE、TFIIF 和 TFIIH 等）的结合。后续研究使用同样的方法表明，一旦起始，Med、TFIID、TFIIH 和 TFIIE 则被留在后面，形成一个有利于重新起始的支架，但是，Pol II、TFIIF 和 TFIIB 离开启动子（Yudkovsky et al. 2000）。固定化模板分析与光交联结合，也可用于研究 GAL4 和 GCN4 激活结构域在酵母提取中的靶标（Fishburn et al. 2005）。

固定化模板分析的一个缺点是，当对复合物进行功能分析或通过免疫印迹法分析时，粗制通用因子制剂与 DNA 和树脂的背景结合会引起很多问题。因此，研究中必须包含仔细设置的特异性对照，包括启动子和激活因子突变体，但不局限于此。此外，可能需要对漂洗条件和缓冲液条件进行系统分析，以优化结合效果。许多此类问题都包含在 Hahn 的研究中，包括对复合物比活性的讨论。有关实验和方法学的其他细节，可查阅 Hahn 实验室网址（http://www. fhcrc. org/science /labs/hahn）。方案 12. 6 也已被证实在哺乳动物系统中在 DNA 和染色质 DNA 模板上都是成功的。

凝胶过滤

凝胶过滤法在概念上与固定化模板法相似，可用于研究转录复合物组装的不同阶段。该方法最初用于研究 Pol III 转录复合物中因子的加成（Carey et al. 1986），简单地说，就是用核提取物或分级分离的因子在质粒 DNA 模板上组装复合物，然后将组装反应加到凝胶过滤柱上（如 Sepharose 4B），通过分子筛选效应将质粒 DNA 及其任何结合因子从游离的、未结合的因子中分离出来。最后，含有质粒 DNA 和结合因子的排除体积用免疫印迹或转录互补分析法分析。

凝胶过滤柱对鉴定转录复合物组装中的早期步骤非常适合，并已成功用于研究激活因子介导的 TFIIB 与 TFIID 的结合。在一例研究中，TFIID 和 TFIIB 在 GAL4-VP16 存在与否的情况下孵育，并通过添入缺失因子和核苷酸补充排除体积馏分，结果表明，只有激活因子存在时，TFIIB 才与 TFIID 结合。令人惊异的是，激活因子促进 TFIIB 的结合，但不促进其他通用转录因子的结合，即使 TBP 取代 TFIID，也是这种情况（Choy and Green 1993；Roberts et al. 1995）。但是，重要的是，要注意诸如 SWI/SNF 的大型复合物往往污染排除体积，即使它们没有稳定结合到 DNA 上。因此，解释结果时必须小心，并且当存在不确定性时，应在无 DNA 的情况下测试纯复合物的迁移。

研究活化作用的高锰酸盐探测

我们已经讨论了使用高锰酸盐技术研究开放复合物的效用。但是，由于开放复合物测量完整的 PIC 形成，该项技术还可用于动力学和热力学实验，以定量测定激活因子刺激的转录复合物（Wang et al. 1992a）。

在一个典型的体外反应中，质粒 DNA 模板在激活因子存在与否的情况下在核提取物中孵育，并加入 ATP 使 DNA 解链。预孵育一段时间，让活化复合物组装，再加入高锰酸钾。终止修饰反应，经修饰的 DNA 在热循环仪上使用 *Taq* 酶线性扩增反应进行引物延伸。如前所述，*Taq* 酶在被修饰的胸腺嘧啶处停止延伸。将活化反应和基础反应

的开放复合物信号在聚丙烯酰胺/尿素凝胶上进行比较（见方案 12.7）。

该方法最直接的应用是确定激活因子是否刺激开放复合物的形成，以及开放复合物的相对水平是否与转录水平相当（Wang et al. 1992a）。当用于确定激活因子对起始与再起始的影响时，这种分析的结果能提供相当多的信息。由于开放复合物是第一轮转录的有效产物的度量，所以开放复合物应首先与单轮分析中产生的转录产物相比较。将模板的转录限制到单轮的最有效的方法是在形成 ATP-依赖性开放复合物之后进行十二烷基肌氨酸钠考验实验。正如以前所讨论的，一定浓度的十二烷基肌氨酸钠允许起始前聚合酶的延伸，但可以有效地阻止再起始。如果激活因子对开放复合物的刺激倍数和单轮分析中转录的刺激倍数（通过光密度计或磷屏成像仪定量确定）相等，则可以暂时推断激活因子在 PIC 形成的步骤中起作用。如果开放复合物的量在激活因子存在与否下相似，但单轮转录分析表现出刺激作用，则可以推断激活因子影响起始后的一个步骤（即启动子逃脱或延伸）。如果再起始速度和起始速度成比例，则在多轮分析中，开放复合物信号与转录信号应仍然成比例。但是，如果开放复合物在有激活因子时被刺激，但刺激倍数少于多轮转录分析中观察到的刺激倍数，则可以暂且推断激活因子刺激再起始的速度大于对起始的刺激速度。这样的观察结果可能成为进一步研究再起始机制的根据（见下文）。

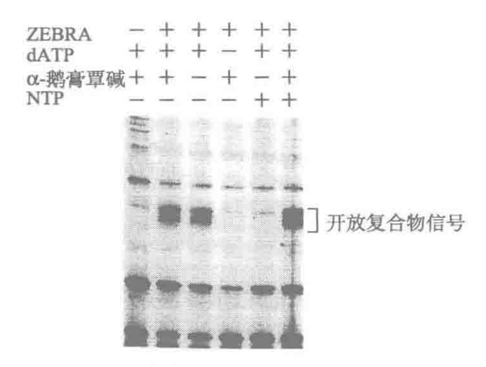


图 12.15　开放复合物形成与转录的比较（详见正文）（经许可转载自 Ellwood et al. 1998, © Spring Harbor Laboratory Press）。

高锰酸盐探测的另一个应用是比较不同模板上的开放复合物和转录信号，目的是研究调控现象，如协同作用。考察这一应用对理解如何利用和解释该技术是有益的。图 12.15 和图 12.16 比较了开放复合物的形成和转录，以研究裂解周期反式激活因子，即 Z Epstein-Bar 复制激活因子（ZEBRA）的激活作用（Chi and Carey 1993）。为此构建了模式报告模板，该模板含有不同数目串联重复的位于腺病毒 E4 核心启动子上游的 ZEBRA 结合位点。围绕 E4 起始位点的区域含有 6 个胸腺嘧啶，因此，该区域是高锰酸盐反应的理想底物。在有 ZEBRA 和 dATP（供能）的情况下，当报告模板在 HeLa 核提取物中孵育时，围绕起始位点的胸腺嘧啶变得对高锰酸盐敏感（括号内区域，图 12.15）。ATP 是复合物形成所必需的，因为省略 dATP 会废止这种效应。开放复合物具有转录活性，并且归因于 Pol II，因为如果加入核苷酸，高锰酸盐敏感性就会消失。这种现象与 Pol II 撤离启动子的延伸和相伴随的起始位点周围 DNA 的重新闭合是一致的。但是，如果将核苷酸和延伸抑制剂 α-鹅膏覃碱一起加入，复合物则留在原地，从而确认它们确实是具有转录能力的 Pol II 开放复合物。

图 12.16 表明，在转录分析中检测到的 ZEBRA 的位点依赖性协同效应（图 12.16，左图）和这些相同模板上形成的开放复合物的相对数量是相对应的（图 12.16，右图）。因此，在多轮分析中，ZEBRA 对转录的协同效应首先表现为转录复合物组装的协同增

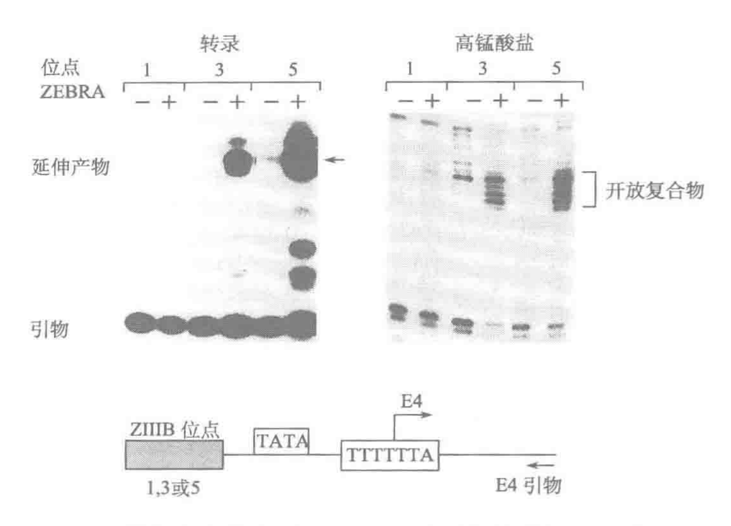


图 12.16 开放复合物形成，与转录相比较（详见正文）（经许可转载自 Ellwood et al. 1998，© Spring Harbor Laboratory Press）。

加。这个例子说明，转录与转录复合物组装有点成比例。假如转录结果与开放复合物结果有分歧，则需研究 ZEBRA 对延伸的影响。

高锰酸盐分析也可用于研究再起始。如果让 PIC 组装，并且之后加入核苷三磷酸，聚合酶会起始并延伸转录。加入核苷酸后再加入 α-鹅膏蕈碱，会使聚合酶延伸在下游停止，但不抑制再起始。然后，可以比较开放复合物再形成和起始开放复合物形成所需要的时间。使用该方法表明，有激活因子的再起始速度比起始速度更快（Jiang and Gralla 1993）。

高锰酸盐分析还可用在重建的系统中，以研究转录复合物组装过程中的限速步骤（Wang et al. 1992b）。用 ZEBRA 和 GAL4-AH 进行的动力学实验表明，在 HeLa 核提取物中组装一个完整水平的开放复合物需要 30～60min。根据独立测量，激活因子与 DNA 的结合及 ATP 依赖性异构化是快速事件。在该领域中其他研究的基础上，推测 30 min 延迟是激活因子组装封闭复合物所需要的时间。激活因子影响的步骤可以通过实验确定，即先将通用转录因子组合物与激活因子一起预孵育 30 min，再加入其余因子，以进行开放复合物组装的动力学分析。如果预孵育已经绕过了缓慢的激活因子应答步骤，那么随着缺失组分和 ATP 的加入，开放复合物会立即形成。如果预孵育还未跨越这一步骤，复合物的组装则需要 30 min。的确，ZEBRA 或 GAL4-AH 与粗制 TFIID 及 TFIIA 馏分的预孵育绕过了一个缓慢的步骤。该分析首次表明激活因子在 TFIIA-TFIID 复合物形成过程中起作用，这个问题下文中还会讨论。动力学方法的一个缺点是不能分析复合物的成分，因为 HeLa 提取物中存在未结合的因子。如果目的是物理性分析转录复合物的结合和转录复合物的成分，就必须采用其他方法。

TFIID-TFIIA 复合物的 EMSA 和 DNase I 足迹分析

用于分析活化转录复合物组装的最直接的方法可能是 EMSA 和 DNase I 足迹法。

纯转录因子蛋白的获得已经使这种分析变得较为简单。Sawadogo 和 Roeder（1985a）发现激活因子 USF 增强了 DNA 上 TFIID 足迹的稳定性，但是从来没有正式表明 TFIID 被募集。使用纯化的 TFIID 和 TFIIA 及 EMSA 和 DNase I 足迹法进行的实验，为募集提供了第一个直接的功能性证据（Lieberman and Berk 1994）。该方法很重要，因为已经发现很多激活因子介入 TFIID-TFIIA 结合步骤，并且酵母生化和遗传实验已经确认了这一步骤在基因调控中的整体重要性。很有可能最终发现许多激活因子都能影响这一步骤（综述，见 Hampsey 1995；Martinez 2002）。

在 EMSA 研究中，使用镁-琼脂糖凝胶，而非聚丙烯酰胺凝胶，是因为 TFIID-TFIIA 复合物的块头很大（图 12.17）（见方案 12.8）。将含有 E4 起始位点和上游 ZEBRA 位点的 DNA 模板与重组 ZEBRA、纯化的 HA 标记 TFIID 及纯化的 TFIIA 一起孵育。之后将形成的复合物在镁-琼脂糖凝胶上电泳分级。TFIID（D）或与 TFIIA 相结合的 TFIID（DA）结合微弱，几乎观察不到迁移的复合物。在这些凝胶中单独的 ZEBRA 轻微地迁移探针，但是却强烈地刺激一个包含 TFIID 和 TFIIA 的复合物（DAZ）的迁移。ZEBRA 对 TBP-TFIIA 复合物的组装几乎没有影响，从而说明增加的募集归因于 TAF（未出示）。在这种情况下，TAF_{II}的确切作用还不清楚，它们可能直接与 ZEBRA 相互作用，也可能稳定一个复合物构象，该复合物允许与 TBP 和 TFIIA 的富有成效的相互作用。

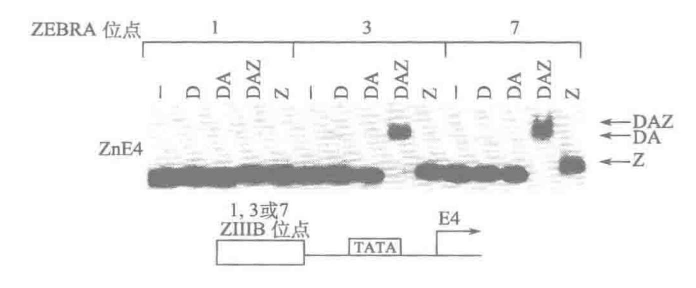


图 12.17　ZEBRA 对协同性 DA 复合物形成的促进作用（详见正文）（经许可转载自 Chi et al. 1995，© Macmillan Magazines Ltd.）。

在 DNase I 足迹分析中也能观察到在 EMSA 中观察到的 ZEBRA 介导的强烈募集，但 EMSA 比较易于定量。Chi 及其同事（1995）使用同样的分析，表明多个 ZEBRA 分子协同募集 DA 复合物，从而表明 ZEBR 与 DA 复合物的不同亚基相接触。TFIIB 与 DA 复合物相互作用的能力需要激活因子，并且 DNase I 足迹分析提示，ZEBRA 能诱导复合物中的构象变化，并且这种变化与 TFIIB 的结合能力有关（Chi and Carey 1996）。因为在没有激活因子时，TFIIB 与 TA 结合得很好，但与 DA 复合物却不能结合，因此推断 TAF 通过封阻 TFIIB 向 TBP 的靠近对转录产生负面影响，暗示激活因子诱导的构象变化能抵消这种抑制作用。

尽管这些结合分析表明 DA 和 DAB 复合物是关键的靶标，但其他方法表明激活因子的加入可以进一步稳定这些复合物。利用相似的模式模板和 ZEBRA 表明，ZEBRA 与 DAB 的预孵育及 USA 共激活因子馏分（包含 Med）能稳定复合物，使它能够耐受

低剂量十二烷基肌氨酸钠的考验而存活下来，并且使转录随着其他因子的加入而能够进行（Lieberman and Berk 1994）。此外，ZDABUSA 复合物的形成也能稳定该复合物，使其能承受 ZEBRA 结合位点寡核苷酸的考验。相似的方法后来被应用于 IFN-β 增强体，结果表明，IFN-β 增强体也刺激含有 DAB 和 USA 馏分的复合物的组装（Kim and Maniatis 1997）。非常可能的是，稳定复合物的 USA 馏分的主要成分是 Med，尽管不能排除其他 USA 组分（如 PC4）的作用。

许多研究表明激活因子除了与 TFIIA、TFIID 和 TFIIB 相互作用外，还与大的共激活因子复合物发生紧密互作。因为这些复合物已被纯化，很可能本章中所述的分析方法将适用于理解这些复合物如何搭建到 TFIID 和 TFIIA 上，以及它们又是如何影响染色质模板上的起始、再起始和转录。

TFIID 亚复合物的组装和分析

生化方法的优势之一是能够控制参与反应的蛋白质组分，这会相应地促进特定假说的提出并检验。虽然 GTF 中有几个是单体或二聚体，但较大的 GTF（如 TFIID 和 TFIIH）带来一个问题：在反应过程中各亚基不容易组装。需要付出很多努力组装并分析这些复合物，最简洁的方法之一是利用纯化的 TAF_{II}亚复合物（表 12.3）来确定激活因子-TAF_{II}相互作用和活化转录之间的直接关系（J. L. Chen et al. 1994）。前提很简单：将 TAF_{II}和激活因子之间的相互作用与功能分析联系起来。尽管该方法已应用于很多激活因子-TAF_{II}相互作用，但我们只提供其中之一用于例证该实验方法。

将嵌合体激活因子 GAL4-NTF-1 用于亲和层析实验，以鉴定哪些果蝇 TAF_{II}结合，发现只有 $TAF_{II}150$（TAF2）紧密结合。然后通过表位标记 $TAF_{II}250$（TAF1）与抗体亲和树脂的结合，构建了 TAF_{II}亚复合物，HA 单克隆抗体连接到蛋白质 A-Sepharose 上。之后将 10 倍过量的 TBP 及其他相关 TAF_{II}与 $TAF_{II}250$ 树脂一起孵育，促进亚复合物的形成，然后用 HA 肽将其从树脂上洗脱。TAF_{II}和 TBP 以一种非常特异的方式

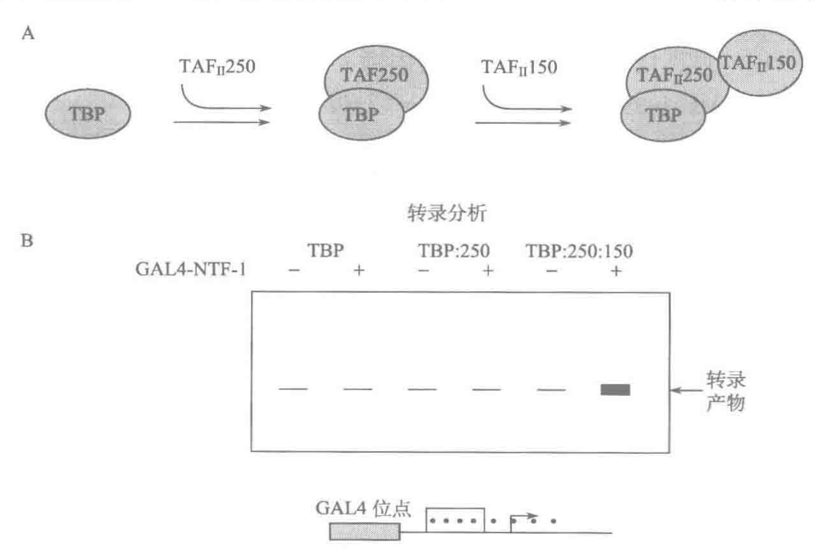


图 12.18 TAF 重建实验。(A) TAF 组装示意图；(B) 体外转录激活。

组装成 TFIID（图 12.18A）。首先 $TAF_{II}250$ 与 TBP 相结合，之后，大多数其他 TAF_{II} 搭建到 $TAF_{II}250$-TBP 复合物周围。$TAF_{II}150$ 与 $TAF_{II}250$-TBP 复合物结合，形成 $TAF_{II}150$-$TAF_{II}250$-TBP 亚复合物，之后才能用体外转录反应检测 TAF_{II}亚复合物支持 GAL4-NTF-1 转录活化的能力（图 15.13B）。在 $TAF_{II}150$-$TAF_{II}250$-TBP 亚复合物存在的条件下，GAL4-NTF-1 被有效地激活，而 TBP 或 TBP-$TAF_{II}250$ 存在时，GAL4-NTF-1 未被激活。因此，激活因子和特定 TAF_{II}的互作能力与其在亚复合物中刺激转录的能力直接相关。

虽然我们强调了 TAF_{II}和激活因子的相互作用，但在基因表达领域中利用亚复合物探测机制变得日益普遍。多种大分子复合物的亚基已经从杆状病毒和其他表达系统（见第 10 章）中克隆并纯化出来，并被重建成各种亚复合物应用于机制研究，如 TFIIH（Tirode et al. 1999）、SWI/SNF（Phelan et al. 1999）和哺乳动物多梳蛋白复合物（Lavigne et al. 2004）。

调解因子亚复合物

现在已知酵母调解因子含有 20 个亚基，它们组装成三个模块，分别称为头部、中部和尾部。另外的一个抑制模块包含一对“细胞周期蛋白依赖性激酶-细胞周期蛋白”和其他两个亚基。这个模块的作用仍然存在争议（综述见 Myer and Young 1998; Björklund and Gustafsson 2005）。Pol II 似乎是一个稍微稳定的 Med 的组分。通过电子显微镜观察研究，Med 复合物从酵母到人类显示出高度的结构保守（Asturias et al. 1999; Naar et al. 2002; Taatjes et al. 2002）。然而，Med 的这些结构上相似的形式并不对应于功能上具有活性的形式，从而表明 Med 的模块性也许是最高程度的保守（Chadick and Asturias 2005）。

已经在整个真核生物中鉴定了 Med 亚基，证明存在有限的但显著的同源性（表 12.4）（Boube et al. 2002）。这些单个亚基是几种转录激活因子蛋白的靶标，这与 Med 作为调控性因子靶标的思想是一致的。例如，在哺乳动物细胞中，Med1 似乎是一个某些核受体（NR）的直接靶标，这些核受体通过配体结合域与 MED1 结合（Ito et al. 2000）。此外，VP16 的激活结构域结合 MED25（Mittler et al. 2003）。这些相互作用可以在有完整的调解因子或其亚基伴随的情况下，用 GST-NR 和 GST-VP16 融合蛋白在体外再现。其中，调解因子亚基由杆状病毒表达系统或通过体外转录/翻译合成。因此，如果对一个特定的激活因子是否与 Med 相互作用感兴趣，该激活因子的 GST 融合蛋白就可用于捕获实验。理想的情况下，从粗提物开始，使用对几种 Med 亚基的免疫印迹，确定 Med 是否保留在柱上。然后，加入依靠体外转录和翻译产生的单个的亚基。如果特定的亚基已被鉴定，则可以用激活因子突变体和 Med 亚基的缺失突变体尝试一个突变分析。如果这个有效，一系列使用染色质免疫沉淀、RNA 干扰和其他方法的体内实验，对证实相互作用的有效性将是必需的。

重要的是要注意，借助对几个不同亚基的标签免疫亲和，从 HeLa 细胞中纯化人类 Med 时，产生了两种形式的复合物（Sato et al. 2004）。其中，一种复合物富含 CDK8 模块（CDK8、CyclinC、MED11、MED12），但相对缺失 Pol II 和 MED26（CRSP70）；另一种复合物富含 Pol II 和 MED26，但缺失 CDK8 模块（Paoletti et al. 2006）。是否

CDK8 形式是一种负作用形式，而 MED26 形式是活性形式，一直存在争论。

目前，已经进行了少数有限的将各种亚基组装成亚复合物或子模块的研究。表达 Med18、MED20 和 MED29 的杆状病毒构体已被用于创建一个头部子模块（Sato et al. 2003）。由于已知 MED29 在果蝇中与 DSXF 相互作用，这种相互作用可能调控性别决定相关基因。此外，MED8、MED18 和 MED20 在酵母中也形成一个头部子模块（Larivière et al. 2006）。用大肠杆菌表达系统内的共表达组装该子模块，发现这个子模块通过 MED8 亚基与 TBP 相互作用。这个研究表明，Med 不仅与激活因子相互作用，也与通用转录因子（如 TFIID）相互作用（Johnson et al. 2004）。目前已经有许多在酵母、哺乳动物细胞和果蝇中已构建的 Med 亚复合物的例子。随着越来越多的实验室研究这些相互作用，一定会出现更加完整的各种激活因子与 GTF 相互作用的图景。人类和酵母的所有 Med 亚基都已被克隆，并且可从各种来源获得这些亚基。

技术

核提取物制备：概述

无细胞提取物作为调控因子的来源，可用于 DNA 结合、体外转录和蛋白质纯化等实验（Ge et al. 1996；Maldonado et al. 1996；Roeder 1996）。历史上，有三种类型的提取物成功地用于体外重现 Pol II 转录。第一类无细胞系统是细胞质 HeLa 细胞 S100 提取物（Weil et al. 1979a），该系统在补充 Pol II 的情况下引起质粒模板的体外转录。紧随其后的是全细胞提取物（Manley et al. 1980），这种系统易于制备，无需加入纯化的 Pol II。最后，还开发了使用 HeLa 细胞的核提取物（Dignam et al. 1983a）。后来，又开发了其他类型的核提取物，这些核提取物制备方法各不相同，并且更适于利用固体组织（Gorski et al. 1986；Stuempfle et al. 1996）。但是，由于它们的转录效率和纯度，Dignam 核提取物获得了最广泛的应用。

虽然已经发展了有效的微量提取物方案，以便利用少至 3×10^7 个细胞的样品（Lee et al. 1988；Osborn et al. 1989），但传统的核提取物制备依赖于大量的细胞（10^{10}或更多）。制备核提取物的各种各样的方法随细胞类型的不同而有所变动。利用平皿、摇床或滚瓶中的组织培养细胞比利用固体组织更容易，但方法的选择主要取决于应用和用于制备提取物的组织或细胞系的可得性。如果研究的目的是获得一个系统，用于为了加入外源组织特异性调控蛋白或用于通用机器的研究，那么按 Dignam 等（1983a）的方法制备的 HeLa 核提取物是通用转录因子的良好来源，并且在比活性（即每毫克提取物的转录）基准上活性很高。一旦设备到位，就可以在旋转培养中以除劳动力外约 8 美元/升的成本产生 HeLa 细胞（而商品来源的细胞需 25 美元/升或更高）。

HeLa 核提取物相对来说易于制备，尽管培养细胞所需的时间（倍增时间 24 h）和设备存在不利之处。在 Dignam 和 Roeder 方法中，细胞在低渗缓冲液中被膨胀并裂解，细胞核通过离心从细胞质提取物中分离出来。然后，产生的细胞核沉淀在中盐缓冲液（0.42 mol/L NaCl）中孵育，以提取转录成分。在该程序中有最小的蛋白质水解，并且

每升培养物产生 1～2 g 细胞（取决于收集时的细胞密度），可产出 1～2 ml 浓度为 6～10 mg/ml 的提取物。这足够用于 50～100 个体外转录反应，反应的个数取决于反应的规模。其他细胞系，包括一些淋巴细胞，也可以在大规模旋转培养装置或滚瓶中培养，但有时这些细胞的生长需要 CO_2（与利用碳酸氢盐和高压灭菌培养液的 HeLa 细胞不同）。尽管如此，从这些细胞中也可制备高质量提取物（Foulds and Hawley 1997）。

如果目的是研究一个基因在其天然细胞类型中的转录，并且难以获得组织培养源，很多研究者就从固体组织中获得核提取物。固体组织一般更难操作，因为它们含有很多细胞类型，包括结缔组织。很大的精力花费在从组织中制备细胞核和防止蛋白酶解。比如，一只大鼠含有 10 g 肝组织，这些肝组织最终只产生 6～7 mg 高质量核提取物，少于从 1 L HeLa 细胞中获得的核提取物。固体组织也必须首先用匀浆器破碎，然后用机械 Dounce 裂解。随后通过蔗糖密度梯度离心纯化细胞核，最后进行抽提。包括核酸酶和蛋白酶在内的污染物的存在，在核提取物制备期间可严重影响提取物的活性，而某些组织（如肝脏和肠）富含这些酶。由于这一原因，某些组织或细胞系比其他组织或细胞系产生更富有活性的提取物。提取条件的优化对于获得通用因子和抑制因子之间最有利的平衡来说通常是必需的，这视所做实验及其误差而定（有关讨论见 Dignam et al. 1983a）。

Dignam 和 Roeder 提取方法（Dignam et al. 1983a）仍然是最受偏爱的大规模制备核提取物的方法，它可靠、简单，且适用于多种组织培养细胞系（方案 12.1）。其他潜在的有用方案包括固体组织核提取物和全细胞提取物制备。Gorski 等（1986）描述了一种从固体组织制备提取物的方法。Stuempfle 等（1996）提供了这一方案的修改，宣称修改方案更容易且应用更广泛（Stuempfle et al. 1996）。最后，当核提取物失败时，一种合理的尝试途径是使用全细胞提取物（Manley et al. 1983）。

方案 12.1 Dignam 和 Roeder 核提取物

对于 Dignam 和 Roeder 核提取物制备（Dignam et al. 1983），首先通过离心收获细胞，洗涤细胞，并将细胞重悬于低渗缓冲液中，低渗缓冲液会导致细胞膨胀。随后用手动 Dounce 匀浆，将细胞裂解。通过离心使细胞核沉淀，弃去细胞质上清液，Dounce 匀浆使核沉淀重悬于中浓度盐溶液中。将重悬物搅拌 30 min，抽提转录因子，再一次离心沉淀细胞核，产生的上清液在缓冲液 D 中进行透析，以用于转录实验。

时线与安排

收获细胞后，Dignam/Roeder 核提取物的制备需要 1 天时间。虽然使用新鲜细胞极其重要，但在制备好细胞核后，可停止该方案，并将核沉淀冷冻，以备后用。此外，可在透析前停止该方案。所有的操作都必须在冰上或 4℃冷柜中进行。缓冲液、离心管、转头以及 Dounce 匀浆器在使用前都要冷却。制备好的提取物于 4℃保存时，半衰期为 12 h，故将透析时间降到最低非常重要。

提纲

从 HeLa 细胞中制备核提取物（1 天）

步骤 1：配制缓冲液并预冷（1 h）。

步骤 2：收集并裂解细胞（2 h 或更长，根据细胞量而定）。

步骤 3：制备核提取物（4 h）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

牛血清白蛋白（BSA）

Bradford 染料浓缩液〈!〉

缓冲液 A

- 10 mmol/L HEPES-KOH（pH 7.9，4℃）〈!〉
- 10 mmol/L KCl
- 1.5 mmol/L $MgCl_2$〈!〉
- 临用前加入：
  - 0.5 mmol/L 二硫苏糖醇（DTT）〈!〉
  - 0.5 mmol/L PMSF〈!〉

缓冲液 C

- 0.2 mmol/L EDTA
- 25% 甘油（V/V）
- 20 mmol/L HEPES-KOH（pH 7.9）〈!〉
- 0.42 mol/L NaCl
- 1.5 mmol/L $MgCl_2$〈!〉
- 临用前加入：
  - 0.5 mmol/L DTT〈!〉
  - 0.5 mmol/L PMSF〈!〉

缓冲液 D

- 0.2 mmol/L EDTA
- 20% 甘油（V/V）
- 20 mmol/L HEPES-KOH（pH 7.9）〈!〉
- 0.1 mol/L KCl〈!〉
- 临用前加入：
  - 1 mmol/L DTT〈!〉
  - 1 mmol/L PMSF〈!〉

干冰（可选；见步骤 3：7）〈!〉

HeLa 细胞

在旋转瓶中 37℃生长 HeLa 细胞，细胞也可从商业来源获得［如国家细胞培养中心（National Cell Culture Center），http://www.nccc.com］。

液氮（可选；见步骤 3：7）〈!〉

磷酸缓冲盐（PBS）（10×原液）

试剂	g/L
KCl〈!〉	2.0
KH_2PO_4	2.0
NaCl	80.0
$NaHPO_4 \cdot 7H_2O$	21.6

用前稀释到 1×。

台盼蓝（可选）〈!〉

设备

离心机（RC-3B，Sorvall）

离心机转头（SS-34，Sorvall）

电导计

离心瓶（1-L）

透析管（约 10 000 M. W. 截留）

血细胞计数板

匀浆器（Dounce，配有 B 型杵）

组织

磁力搅拌器和搅拌棒

显微镜

吸液管

转瓶

试管（用于离心，标有刻度）

程序

步骤 1：准备缓冲液和冷却设备

使用前，准备并在冰上或 4℃冷柜中冷却所有缓冲液、离心管、转头及 Dounce 匀浆器。

步骤 2：收获和裂解细胞

1. 收获之前，使用血细胞计数板对 HeLa 细胞计数。通过 3000 r/min 离心 10 min 收获细胞，并使细胞密度为（0.5×10^6～1.0×10^6）个细胞/ml。

理想情况下，大于 2 L 的细胞培养物，用 Sorvall RC-3B 离心机在 1 L 的离心瓶中收获。

2. 每升培养物的细胞沉淀块用 10 ml PBS 轻轻地洗涤。3000 r/min 离心 10 min，弃上清，加入等体积 PBS 悬浮细胞。

3. 将细胞悬液转移到锥形刻度离心管中。3000 r/min 离心 10 min，弃上清。所有后续步骤在 4℃操作。

4. 使用离心管上的刻度，确定细胞沉淀的体积。细胞沉淀重悬在 5 倍体积的缓冲液 A 中。理想情况下，1 L 培养细胞产生的 1～2ml 细胞沉淀。

5. 转移上清到一个配 B 型杵的 Dounce 匀浆器中，击打 10 次裂解细胞。该步操作要轻缓，尤其是 B 型杵下推要缓慢温和，以防止核裂解。如果悬液变得黏稠，即停止匀浆。

6. 显微镜下检查裂解物，观察细胞裂解情况。

也可以用台盼蓝染细胞监测裂解。

步骤 3：制备核提取物

1. 将步骤 2：5 获得的裂解物在 Sorvall SS-34 转头中 10 000 r/min 离心 20 min。弃去含有细胞质物质的上清，用吸液管或纸巾小心除去离心管管口附近的脂质。

2. 通过柔和的 Dounce 匀浆（即用 B 型杵击打 10 次）将粗细胞核沉淀重悬于缓冲液 C（3 ml/10^9 个细胞）中。

3. 用磁力搅拌棒于 4℃下将细胞核悬液轻轻搅拌 30～60 min。

4. 用 Sorvall SS-34 转头将悬液 10 000 r/min 离心 30 min。

这一步从细胞核中抽提转录组分。

5. 将清亮的上清转移至透析袋中，在 50 倍体积缓冲液 D 中于 4℃下透析 5 h。

必须剧烈旋转透析袋，以在所述时间内实现彻底透析。较长时间的透析会使转录活性下降。测定电导率，确保透析彻底。

6. 透析物用 Sorvall SS-34 转头 4℃下 10 000 r/min 离心 20 min，弃沉淀。

7. 上清为细胞核提取物，分装成 0.1～1 ml 小等份，在液氮中或干冰上速冻，储存于－80℃。

8. 用 Bradford 法分析一小份细胞核提取物，同时以 BSA 作为标准物（见疑难解答），分析核提取物的蛋白质浓度。通常，蛋白质浓度为 6～10 mg/ml，或从 10^9 个细胞可获得 15～20 mg 蛋白质。

附加说明

1. *抽提步骤*。抽提步骤中的盐浓度对于 HeLa 提取物应该是最适的。较高的盐浓度会导致抑制物开始抽提，而较低的盐浓度不能有效地抽提通用转录组分。其他细胞类型可能会表现出不同的最佳抽提条件，如 Dignam 等（1980）所述，这些最佳抽提条件应该依据经验而定。

2. *提取物浓度*。保证提取物有活性的关键是蛋白质浓度高。浓度低于 5 mg/ml 的提取物通常活性微弱。有些研究人员将 Dignam 和 Roeder 程序与硫酸铵浓缩步骤联用，将固体硫酸铵（0.3～0.4 g/ml 上清）加到最终的提取物中（步骤 3：4 后），轻轻搅

拌。这一浓度的硫酸铵足以使提取物中的绝大多数蛋白质发生盐析或沉淀。30 min 后，离心收集沉淀，并重悬于缓冲液 D（比起始体积浓缩 2～4 倍）中，透析除去过量的硫酸铵。

3. *透析时间*。当对提取物进行透析时应考虑两点。第一，较长的透析时间会导致过量的蛋白质沉淀，使蛋白质的浓度和提取物的活性下降。通常透析时沉淀的蛋白质不可能被重悬。第二，提取物经受随机的蛋白酶降解且蛋白质失活。因此，加入新鲜的蛋白酶抑制剂和还原剂对于维持提取物的活性是必需的。在这两种情况下，重要的是通过透析液的剧烈搅拌、大孔径透析管（约 10 000 M. W. 截留）的使用、表面积的优化（即通过使用细长透析管与粗短透析管进行比较），尽量缩短透析时间。保存前要测定透析物的电导率，确定透析物已经和缓冲液相平衡。如果采用了上述策略，提取物通常可以在 5 h 或更短的时间内透析至平衡。

4. *微量提取物*。对于有些实验，或对于优化抽提条件，有必要从少量细胞中获得提取物。有两项技术能用 3×10^7 个细胞（即相当于 60 ml HeLa 细胞）制备提取物。第一种技术是，使细胞流经注射针而使其裂解，再离心沉淀细胞核，在微量离心管中抽提（Lee et al. 1988）。第二种技术是，用低渗缓冲液和 NP-40 裂解细胞（Osborn et al. 1989）。这些技术产生的提取物可用于 DNase I 足迹分析、EMSA 和体外转录实验。

疑难解答

无转录活性

可能原因：提取物有问题。

解决办法：首先利用对照提取物检测所用技术。Promega、Protein One 及其他生产厂家都出售具有转录活性的 HeLa 提取物。如果商业化提取物对实验模板起作用，应考虑提取物中的蛋白质浓度可能太低，或盐浓度可能太高。通过 Bradford 方法检查蛋白质浓度，测定电导率确定盐浓度，确定它们是否在可接受的范围内（见方案 12.2）。

可能原因：提取物中含有抑制物。

解决办法：进行混合实验。将提取物混合进对照提取物中，以确定其是否具有抑制作用。

可能原因：由于 HeLa 细胞的质量或其他未知原因，提取物制剂偶尔会失败。

解决办法：对这种提取物可偶尔补充极限因子，以复苏它们，但是最好的建议是以新鲜的缓冲液和细胞重新开始。

方案 12.2　使用 HeLa 细胞提取物和引物延伸的体外转录

关于在无细胞系统中重现外源基因转录的研究，最初是为 Pol III 进行的。RNA Pol III 转录的基因（如 5S 和 tRNA）一般较小（50～150 bp），在蟾蜍（*Xenopus*）卵母细胞提取物（Birkenmeier et al. 1975）或 HeLa 细胞提取物（Weil et al. 1979）中加入核苷酸和外源质粒 DNA 模板时，能合成这些基因的全长转录本。许多用于研究 Pol

II 转录的方法都来源于早期的 Pol III 研究。虽然已经创立了能正确转录外源模板的 Pol II 系统（Weil et al. 1979；Manley et al. 1950），但这些系统的效率都不及 Pol III，而且就绝大部分而言，还不能产生全长 mRNA 转录本。然而，这些系统却为随后的转录机器的分级分离和表征充当重要的出发点（Matsui et al. 1980；Segall et al. 1950）。最初的研究采用了线性质粒模板的失控转录。为了表明体外系统生成具有天然体内 5′端的 RNA，通过 RNase 指纹定位 5′端是早期研究的重要方面。最早的研究利用腺病毒和 SV40 基因作为 DNA 模板，因为已知这些病毒含有强组成型启动子，且存在初步的病毒遗传学，能帮助阐明其机制。

虽然多年来基本方法没有多大改变，但是诸如引物延伸和无 G 序列盒等的分析方法已经取代了失控转录和指纹法，以测定转录和定位 mRNA 5′端。在许多体外研究中，由于核提取物的效率的提高，它已成为一种首选提取物（Dignam et al. 1983）。随着该领域的研究重点从理解转录机制的基本方面（即通用转录因子功能）向理解调控的转移，天然细胞模板已经开始逐渐取代病毒基因。

时线与安排

通常来讲，可以将带有激活因子结合位点的 DNA 模板和提供通用转录因子及相应激活因子的核提取物一起孵育，以合成 mRNA。在某些情况下，激活因子存在于提取物中（如 Spl），此时就要对缺失或含有位点的模板进行比较。之后，通过引物延伸分析对 mRNA 进行定量。在引物延伸分析中，首先将标记的短 DNA 寡核苷酸和 mRNA 的一个互补区域退火，之后通过反转录酶将引物延伸到 mRNA 的 5′端，然后在聚丙烯酰胺/尿素凝胶上分级分离并检测所生成的标记 cDNA。产物的量是对激活因子引起的转录活化的一种度量。此外，也可以用失控转录或无 G 序列盒分析代替引物延伸分析。模板通常是裸 DNA，但染色质重建系统正日益受欢迎，许多研究人员都在使用染色质模板。

任何研究的最初目的都是从合适的细胞类型中制备足够质量的提取物，并通过改变多个反应参数优化模板的转录。有了提取物、模板及试剂之后，最初的体外转录反应和产物分析可在两天内完成。关于提取物制备的详细方法，见方案 12.1。第二天的引物延伸反应需要 ^{32}P 末端标记的引物。第 3 章（方案 3.1）含有通过多聚核苷酸激酶进行末端标记的方案。

第 1 天：体外转录和引物标记

第 2 天：反转录和 PAGE 分析

提纲

体外转录（4 h）

步骤 1：准备缓冲液（30 min）。

步骤 2：体外转录反应和产物的分离（2 h，30 min）。

引物杂交、延伸和凝胶电泳（8 h）

步骤 1：准备缓冲液（30 min）。

步骤 2：用预先配制的丙烯酰胺/尿素混合液倒制 10%尿素凝胶（30 min）。

步骤 3：引物杂交和延伸/反转录反应（4 h）。

步骤 4：反应产物的 PAGE 分析（4 h）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

激活因子蛋白（200 ng/μl）

乙酸铵（1 mol/L）〈!〉

缓冲液 D

　0.2 mmol/L EDTA

　20% 甘油（V/V）

　20 mmol/L HEPES-KOH（pH 7.9）〈!〉

　0.1 mol/L KCl〈!〉

　临用前，加入：

　　1 mmol/L DTT〈!〉

　　1 mmol/L PMSF〈!〉

载体 DNA（50 ng/μl）

　非特异性 DNA 最好是用于克隆被检测启动子的质粒载体（如 GEM3 和 pUC18）。

氯仿〈!〉

DNA 模板

干冰〈!〉

乙醇（70% 和 100%）〈!〉

HeLa 细胞核提取物（见方案 12.1）

杂交缓冲液（2×）

　2 mmol/L EDTA

　600 mmol/L NaCl

　20 mmol/L Tris-HCl（pH 7.6）

　临用前，加入

　　0.2% SDS〈!〉

异丙醇〈!〉

$MgCl_2$（0.1 mol/L）〈!〉

NTP（核苷三磷酸）（每种 25 mmol/L）

苯酚〈!〉

引物（^{32}P-末端标记；0.1 pmol/μl）〈!〉

　对于引物延伸，引物应该有 75℃或 75℃以上的熔解温度，且最好富含 GC，没有任何明显的二分体（防止发夹形成），并定位于距假定起点 60～90 个核苷酸的范

围内。

引物延伸缓冲液

试剂	原液	体积
放线菌素 D〈!〉	1 mg/ml	2 μl
dNTP	25 m mol/L	0.8 μl
DTT〈!〉	1 mol/L	0.2 μl
反转录酶（Promega M1701）	200 U/μl	0.3 μl
反转录酶缓冲液（Promega 1701）	10×	2 μl
RNasin		0.3 μl
ddH_2O		补至 10 μl

准备成用于多个反应的一批混合物。

蛋白酶 K（10 mg/ml）〈!〉

反转录缓冲液（10×）

0.75 mol/L KCl〈!〉

0.1 mol/L $MgCl_2$〈!〉

0.5 mol/L Tris-HCl（在 25℃下 pH 为 8.6）

终止缓冲液

10 mmol/L EDTA

0.2% SDS〈!〉

0.3 mol/L 乙酸钠

50 μg/ml 酵母 tRNA

Tris-HCl（10 mmol/L，pH 8.3）

聚丙烯酰胺/尿素凝胶电泳所需其他试剂

设备

滤纸（如 Whatman 3MM）

盖革计数器

微量离心机

巴氏移液管

磷屏成像仪暗盒（可选；见步骤 4：4）

辐射屏蔽

微量离心管（1.5 ml）

真空干燥机

涡旋振荡器

水浴，设置在 45℃

X 光片（可选；见步骤 4：4）

聚丙烯酰胺/尿素凝胶电泳所需其他设备

程序

体外转录

步骤 1：准备缓冲液

准备缓冲液 D 和终止缓冲液。

步骤 2：转录反应

1. 在一个离心管中，冰上混合：

HeLa 细胞核提取物	15 μl
缓冲液 D	10 μl
25 mmol/L NTP	0.8 μl
DNA 模板	50 ng
载体 DNA	200 ng
0.1 mol/L $MgCl_2$	3.0 μl
激活因子蛋白	1.0 μl
H_2O	补加至 40 μl

HeLa 提取物和缓冲液 D 的准确用量可以有所变动，但总体积应保持 25 μl。同样，特异性和非特异性 DNA 的用量应进行滴定，但总量应保持 250 ng。如果需要，对激活因子进行系列稀释，以确定其最佳浓度。

2. 将反应混合物在 30℃孵育 1 h。
3. 加入 100 μl 终止液和 1 μl 110 mg/ml 蛋白酶 K，55℃孵育 15 min。
4. 加入 140 μl 酚，充分涡旋，离心 2 min，之后转移上清到一新管。
5. 上清中加入 140 μl 酚：氯仿（1：1），充分涡旋，离心 2 min。
6. 转移上清至新管中。为避免有机相污染，仅吸取 130 μl 上清。加入 260 μl 乙醇，充分涡旋，在干冰上静置 10 min。
7. 14 000 r/min 离心 15 min。
   此时应能看到含有 mRNA 的沉淀。
8. 用巴斯德吸管（灼烧移液管吸头，再将之拉长）小心吸出上清。
9. －20℃保存沉淀，或用于引物延伸分析。

引物杂交及延伸

步骤 1：准备缓冲液

准备杂交缓冲液、10×反转录缓冲液和引物延伸缓冲液。

步骤 2：用预先配制的丙烯酰胺/尿素混合液倒制 10%尿素凝胶

在开始实验前约 2 h，倒制 10%聚丙烯酰胺/尿素凝胶，凝胶的聚合需要 30 min。载样前预电泳 30～60 min。关于混合及倒制聚丙烯酰胺凝胶的操作指南，见 Sambrook 和 Russell（2001；pp. 12. 74-12. 81）。

步骤 3：引物杂交和延伸/反转录反应

1. 将体外转录反应（步骤 2：9）得到的沉淀重悬于 9 μl ddH_2O 中。

2. 加入由 10 μl 2×杂交缓冲液和 1 μl 末端标记引物（0.1 pmol）构成的 11 μl 混合液。辐射屏蔽后 37℃孵育 2 h。

   最佳温度应取决于引物，依靠经验确定（见方案 3.1）。此时，引物与 mRNA 发生退火。以下步骤纯化用于引物延伸的 mRNA：DNA 杂交体。

3. 加入 200 μl 1 mol/L 乙酸铵和 200 μl 异丙醇。涡旋，室温静置 10 min。

4. 14 000 r/min 离心 15 min。以拉长的巴斯德吸管除去上清。

   应能看到沉淀，但是沉淀小且疏松。重要的是从此步以后用盖革计数器监测沉淀。

5. 加入 400 μl 冰冷的 70%乙醇，洗涤并简单离心。

   这一步骤洗去残留的乙酸铵。

6. 吸出上清，使沉淀风干，残留的乙醇会抑制引物延伸反应。

7. 将沉淀重悬于 10 μl 10 mmol/L Tris-HCl（pH 8.3）中。在冰上，加入 10 μl 引物延伸缓冲液并充分混匀。

8. 将管迅速转移至 45℃水浴中，孵育 1 h。

9. 置冰上冷却。

步骤 4：反应产物的 PAGE 分析

1. 向反应中加入 20 μl 甲酰胺染色混合液，在 90℃下热处理 3 min，迅速置于冰上冷却。

2. 将 10 μl 样品加到 10%聚丙烯酰胺/尿素测序胶（来自步骤 2）上，跑胶，直至溴酚蓝染料迁移出 2/3 凝胶距离（见疑难解答）。

3. 将凝胶转移到滤纸上，真空干燥凝胶。

   此时，可用甲醇/乙酸混合液固定凝胶，但这不是必需的。

4. 将凝胶在 X 光片上或磷屏成像仪暗盒中曝光（见疑难解答）。

附加说明

1. 通过用限制性内切核酸酶将 DNA 模板线性化并加入放射性标记的 NTP（方案 12.3），可使该方案适用于失控转录分析。一般来说，在距转录起始位点 300 个核苷酸处将 DNA 模板线性化。由于 DNA 末端结合蛋白对提取物的污染，距离起始点更近的线性化常具有抑制作用；由于 Pol II 造成的提前熄火，距离起始位点更远的线性化会减小观察到全长转录本的可能性。

2. 对于引物延伸，引物应有 75℃或更高的熔解温度，并且最好富含 GC，缺乏明显的二分体（防止发夹形成），以及应定位于距假定起始位点 60～90 核苷酸处。

3. 通常，研究者通过滴定反应体系中提取物和 DNA 的用量开始体外研究。缓冲液的 pH，以及 KCl、$MgCl_2$ 的浓度不太重要，但这些参数应落在可接受的范围内。要注意的是，在酵母体外研究中，可以采用有机盐（如乙酸钾和谷氨酸盐）。目前某些研究人员往哺乳动物提取物中加入这些盐和体积排阻剂（如 PVP），以提高这些提取物的效率。

4. 在当前的哺乳动物体外系统中，除非转录过程中有强激活因子或染色质模板参与，否则很少能观察到长距离调控（如通过增强子的调控）。

5. 在研究的最早期阶段，应该始终以每毫升 2 μg 的 α-鹅膏蕈碱设置对照反应，以确保所观察到的转录是由 Pol II 引起的。

疑难解答

没有转录

可能原因：提取物可能不具有转录活性。

解决办法：购买商业化对照提取物和模板（如从 Promega 和 Protein One），以确保方法和技术适当。也可购买 HeLa 细胞，以确保制备提取物的技术是有效的。大量细胞可以从 NIH 赞助的国家细胞培养中心（National Cell Culture Center）（www.nccc.com）获得。

微弱信号

可能原因：条件不是最佳。

解决办法：通过改变 DNA 和提取物进行优化。检查引物延伸反应、引物的比活性等。

可能原因：RNA 中的二级结构。是否观察到比预期的分子质量更小的一系列条带？有时，这是由于倒扑合成或提前暂停造成的。也可能缓冲液已经变质。

解决办法：为了防止前一种情况，尽量提高杂交和延伸温度。有时只需简单地改变引物的位置。对于后一种情况，要注意的是禽源反转录酶对 pH 有些敏感，应更换缓冲液，试用新鲜的或新的 dNTP。如果信号的确弱，检查是否将引物标记到高比活性。另外，对于在体内以低水平转录的基因，不要期望有强信号。病毒启动子（如 AdMLP）比许多天然细胞启动子的活性高出 100 倍。

拖尾凝胶

可能原因：这可能归因于未从转录反应中有效地去除蛋白质，这些蛋白质被带入引物延伸。

解决办法：加入新鲜的蛋白酶 K，使反应脱蛋白。有必要进行一次不加氯仿的酚抽提，以有效去除转录反应中的蛋白质。

方案 12.3 使用 HeLa 细胞核提取物的无 G 序列盒体外转录

发展无 G 序列盒分析是为了简化对环状 DNA 模板上 RNA Pol II 转录的测量（Swadogo and Roeder 1985）。无 G 序列盒是非模板链上缺少 G 残基的 365 个核苷酸的 DNA 片段。原则上，在缺失 GTP 的体外反应中，能够产生全长转录本，而 GTP 的省略导致遍及质粒的大多数随机的、非特异性转录的抑制。这种方法和失控转录一样，能产生放射性标记的 RNA 产物，而直接绕过进行引物延伸或其他间接的 mRNA 测量所需要的必需品和额外时间。与失控转录不同的是，失控转录需要切割过的末端，而无 G 序列盒分析可以在环形超螺旋质粒上进行，超螺旋质粒在许多系统中都是较为有效的模板。实际上，多数粗制系统（如 HeLa 核提取物）都含有少量污染性 GTP，这些 GTP 能够引起少量的背景转录，有时也能引起通读无 G 序列盒的随机上游转录。

为了减少这些人造品，反应体系中通常加入 3′-*O*-Me-GTP。3′-*O*-Me-GTP 是一种 GTP 的链终止类似物，这种类似物当掺入生长的转录本中时，能导致转录终止，很像 DNA 测序中使用的双脱氧核苷酸。反应产物用 T1 RNase 切割，该酶在 G 残基处切割 RNA，从而进一步降低背景转录；无 G mRNA 保持完整，而小的随机 RNA 被消化。

时线与安排

与以前的方案相比，用无 G 序列盒作为模板用于体外转录绕过了对引物延伸步骤的需要。新生转录本直接被放射性 ATP、CTP 或 UTP 标记。可直接通过 PAGE 和放射自显影对反应产物进行分析。反应可以在一天内完成，次日就可以检查数据。最初的无 G 序列盒质粒产生 365 个核苷酸的产物。主要延迟在于将无 G 序列盒克隆在天然转录起始位点的直接下游。有了这种模板，该方案花费一天时间，并涉及将 DNA 模板、核提取物和 [^{32}P] UTP、CTP、ATP 一起孵育。之后，终止反应，通过 PAGE 分离和分级产物。通常，凝胶需放射自显影过夜。

提纲

无 G 序列盒体外转录（6～7 h）

步骤 1：准备缓冲液（30 min）。

步骤 2：准备 6%聚丙烯酰胺/尿素凝胶（30 min）。

步骤 3：体外转录反应（2 h）。

步骤 4：凝胶电泳和产物分析（3～4 h）。

材料

注意：以〈!〉标记的材料的恰当处理见附件。

试剂

乙酸（可选）〈!〉

激活因子蛋白（200 ng/μl）
[α-^{32}P] UTP（10 μCi/μl）〈!〉
缓冲液 D
　　0.2 mmol/L EDTA
　　20% 甘油（V/V）
　　20 mmol/L HEPES-KOH（pH 7.9）〈!〉
　　0.1 mol/L KCl〈!〉
　　临用前加入：
　　　1 mmol/L DTT〈!〉
　　　1 mmol/L PMSF〈!〉
载体 DNA（50 ng/μl）
氯仿〈!〉
DNA 模板
干冰〈!〉
乙醇〈!〉
HeLa 细胞核提取物（见方案 12.1）
甲醇（可选）〈!〉
$MgCl_2$（0.1 mol/L）〈!〉
NTP（核苷三磷酸；只含 ATP、CTP 和 UTP）
苯酚〈!〉
蛋白酶 K（10 mg/ml）〈!〉
RNase T1（1000 U/μl）
　　使用前 1∶50 稀释原液。
终止缓冲液
　　10 mmol/L EDTA
　　0.2% SDS〈!〉
　　0.3 mol/L 乙酸钠
　　50 ng/μl tRNA
　　临用前加入：
　　　10 μg/ml 蛋白酶 K〈!〉
聚丙烯酰胺/尿素凝胶电泳所需其他试剂

设备

滤纸（如 Whatman 3MM）
微量离心机
巴斯德吸管
磷屏成像仪暗盒（可选）
微量离心管（1.5 ml）

真空干燥机
振荡器
X 光片（可选）
聚丙烯酰胺/尿素凝胶电泳所需其他设备

程序

步骤 1：准备缓冲液

准备缓冲液 D 和终止缓冲液。

步骤 2：准备 6%聚丙烯酰胺/尿素凝胶

在开始实验前约 2 h，倒制 6%聚丙烯酰胺/尿素凝胶，凝胶的聚合需要 30min。载样前预电泳 30～60 min。关于混合和倒制丙烯酰胺凝胶的操作指南，见 Sambrook 和 Russell（2001；pp. 12. 74-12. 80）。

步骤 3：体外转录反应

1. 在一个离心管中，混合：

HeLa 细胞核提取物	15 μl
缓冲液 D	10 μl
ATP	250 μmol/L
CTP	250 μmol/L
UTP	10 μmol/L
RNase T1	0. 5 μl
[α-^{32}P] UTP	0. 4 μl
DNA 模板	50 ng
载体 DNA（50 ng/μl）	200 ng
0. 1 mol/L $MgCl_2$	3. 0 μl
激活因子蛋白（200 ng/μl）	1. 0 μl

HeLa 提取物和缓冲液 D 的确切用量可以有所变动，但总体积应保持 25 μl。如果需要，对激活因子进行系列稀释，以确定其最佳浓度。

2. 将反应物于 30℃孵育 1 h。
3. 加入 100 μl 终止缓冲液和 1 μl 10 mg/ml 蛋白酶 K，55℃孵育 15 min.
4. 加入 140 μl 酚，充分涡旋，离心 2 min。转移上清到一新管。
5. 向上清中加入 140 μl 酚∶氯仿（1∶1），充分涡旋，离心 2 min。
6. 转移上清至新管中。为避免有机相的污染，仅吸取 130 μl 上清。加入 260 μl 乙醇，充分涡旋，在干冰上静置 10 min。
7. 14 000 r/min 离心 15min。能够看到包含 mRNA 的沉淀。
8. 用巴斯德吸管（灼烧移液管吸头，再将之拉长）小心吸出上清，风干沉淀。

步骤 4：凝胶电泳和产物分析

1. 将沉淀重悬在 20 μl 甲酰胺染色混合液中。将样品在 90℃下热处理 3 min，之后迅速置于冰上冷却。

2. 将样品平衡至室温。取一半样品加到 6%聚丙烯酰胺/尿素凝胶上（来自步骤 1），电泳至溴酚蓝染料向下迁移到凝胶 2/3 处。

3. 将凝胶小心转移到滤纸上。

　有些研究人员切掉含有未掺入［^{32}P］UTP 的底部凝胶。当凝胶还附着在玻璃板上时也可以用 10%甲醇/5 %乙酸浸泡加以固定。

4. 在真空下使凝胶干燥。

5. 将凝胶在 X 光片上或磷屏成像仪暗盒中曝光（见疑难解答）。

附加说明

1. 无 G 序列盒必须克隆在准确的转录起始位点处，并取代天然下游序列。但是，如果下游元件对启动子活性重要，则它们被无 G 序列盒取代可能导致转录水平下降（Burke and Kadonaga 1997）。

2. 无 G 序列盒转录的信号取决于^{32}P 标记的核苷酸的比活性。即使使用高比活性的核苷酸，依然必须以非标记核苷酸将它们稀释至极限 K_m，以使聚合酶能够起始并延伸转录。例如，以非标记的 UTP 将 500 Ci/mmol ^{32}P 已稀释约 20 倍。通过加入更多的冷 UTP 进行进一步稀释会导致更低的比活性。在上述反应中，往含有 4×10^{-7} mmol 非标记 UTP 的反应中加入 4 μCi 含 4.4×10^{-10} mmol 的 UTP。因此，1000 个 UTP 分子中大约有 1 个 UTP 具有放射性。假定在一个无 G 序列盒中含有等量（约 122）的 A、C 和 U，则在 8 个转录本中只有 1 个转录本真正含有放射性标记的核苷酸。还要注意粗提取物有时被微摩尔量的 NTP 污染，这会进一步限制信号的灵敏度。

3. 无 G 序列盒的大小在某些构体中会有所变动，并且可用较小的无 G 序列盒构建对照模板，该对照模板的产物可以作为内对照同时进行检测。

疑难解答

微弱信号

可能原因：试剂可能需要调整。

解决办法：检查核苷酸的比活性，并作相应的调整。有时需要通过滴定，必须将^{32}P标记核苷酸调至最大量，将非标记的载体调至最小量，以优化信号强度。

没有信号

可能原因：提取物可能没有活性。

解决办法：提取物对对照模板具有转录活性吗？引物延伸法作为一种 RNA 分析方法更有效吗？

观察到信号，但无转录活化

可能原因：有时，在粗制系统中，大部分的信号归因于通读转录，这种通读转录是由强烈的随机上游强启动子及提取物中存在高度污染的 GTP 造成的。

解决办法：通过省略 RNase T1 处理，可区分正常大小的 365 核苷酸转录本和上游通读转录产物。但不幸的是，正确起始的转录本可能由于相同的原因而不能正确地终止。因此，最好的解决办法是通过透析提取物进一步去除污染性 GTP 和加入的 3′-*O*-Me-GTP，尽量减少污染性 GTP。另外，在已知转录活化发生的情况下，重新试用其他测定方法。

共激活因子的纯化（方案 12.4 和方案 12.5）

Roeder 及其同事首次表明，能够转录 mRNA 模板的无细胞系统可通过层析细分为不同的馏分，这些馏分本身并不具有活性，但当混合到一起时，它们在生化上彼此互补，引起转录（Matsui et al. 1980）。分级分离的体外系统的主要优点是，它能使研究人员确定来自特定启动子的转录的限速步骤，并确定激活因子如何影响转录复合物的组装。

三种系统在使用中，包括：①因子缺失的系统（Nakajima et al. 1988，Ozer et al. 1994）；②部分分级分离的系统（见 Wang et al. 1992）；③同质的系统（见 Dignam et al. 1983b；Ge et al. 1996；Maldonado et al. 1996；Tantin et al. 1996；Guermah et al. 2006）。在因子缺失的系统中，从提取物中去除一个或多个因子，或简单地使其失活（即通过免疫或亲和耗竭，或通过某些处理进行层析去除选择性失活某个因子），这种系统的优点是它们易于产生，具有高活性，并紧盯单个因子，可仔细地控制该因子的缺失和随后被加回提取物中；其缺点是提取物的粗制特性和只能单个组分进行操控的有限能力。

在部分分级分离的系统中，HeLa 提取物被进一步细分成含有通用因子的不同子集的馏分。在这些系统中，所有的馏分都可以是粗制的，或者有的馏分是粗制的，而其他馏分可用纯化的天然或重组因子代替。后一种方法产生在转录上有效而生化上可操作的系统，但这些系统仍含有能妨碍结果解释的粗制馏分。

同质的系统含有所有纯的通用因子（TFIIA、TFIIB、TFIID、TFIIE、TFIIF 和 TFIIH），并且对于基础转录具有活性，但它们对激活因子大多不应答，除非补充诸如调解因子的共激活因子。即使这样，这些系统也不总是完全再现用提取物所获得的高水平活性。

本节还描述了纯化某些关键共激活因子使其达到同质的方案（方案 12.4 和方案 12.5），以便这些因子可用于研究激活因子如何影响转录复合物组装的高级分析，包括镁-琼脂糖 EMSA、高锰酸盐探测、DNase I 足迹分析和固定化模板等（见第 11 章）。

方案 12.4 表位标记 TFIID 的纯化

TFIID 是转录复合物装配中最关键的因子之一，这是因为它识别核心启动子，并和染色质及激活因子发生相互作用。Roeder 及其同事发展了从 HeLa 细胞提取物中纯

化 TFIID 的传统方法（Nakajima et al. 1988）。虽然这种方法产生能用于 DNase I 足迹和体外转录分析的、相对较纯的 TFIID，但是该方法需要大量的起始核提取物。人 TBP 的克隆使得 Berk 及其同事利用一种称为表位标签的方法修饰 TFIID，因而在一个简单的两步程序中利用免疫亲和层析就能纯化 TFIID，且纯化的 TFIID 活性损失有限。Berk 将含有流感病毒血凝素（HA）标签的短肽融合到 TBP 的氨基端，并利用反转录病毒转移系统产生了一个稳定表达 HA-标记 TBP 的 HeLa 细胞系（Zhou et al. 1992）。该细胞系的提取物中含有稳定掺入表位标记 TBP 的 TFIID。可以利用磷酸纤维素亲和层析将 TFIID 从这些提取物中部分纯化，然后利用含有与 12CA5 或 HA. 11 单克隆抗体（抗流感病毒表位）交联的蛋白 A-琼脂糖凝胶微珠树脂进行免疫纯化。利用 HA 肽可以将 TFIID 以纯蛋白的形式从免疫亲和树脂上洗脱下来。获得的 TFIID 含有全部的 TAF，能够用于转录、EMSA 和足迹分析，其纯度非常适合于很多其他研究。Roeder 及其同事通过往 TBP 上加 FLAG 标签，发展了一个相似的程序（Chiang et al. 1993）。

时线与安排

第 1 天：准备缓冲液、P-11 及免疫亲和树脂。
第 2 天：在 P-11 上分级分离核提取物。
第 3 天：免疫印迹 P-11 馏分，并将峰馏分与免疫亲和树脂结合过夜。
第 4 天：从免疫亲和树脂上洗脱 TFIID。

提纲

步骤 1：从表达 HA-TBP 的细胞系中制备核提取物（见方案 12.1）。
步骤 2：准备缓冲液、P-11 和免疫亲和树脂（1 天）。

制备好含有 HA-TBP 的核提取物后，整个程序需要 4 天时间。文献 Zhou 等（1992）中描述了细胞生长和提取物制备。表达 HA-TBP 的细胞可从 Arnold Berk 实验室（UCLA）获得。

- 准备缓冲液（1 h）。
- 准备 P-11 柱（2 h）。
- 准备 HA-抗体树脂（4～5 h）。

步骤 3：纯化表位标记的 TBP（3 天）。

- 从核提取物中分级分离 TFIID。
- 使 TFIID 和抗体亲和树脂结合，并洗脱 TFIID。

步骤 4：通过 SDS-PAGE 或免疫印迹法分析纯化的 TFIID。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

抗体（抗-HA，单克隆，12CA5 或 HA. 11；1 mg/ml）

Covance 研究产品（Covance Research Products）销售 HA.11 抗体肽（MMS-101P）；表达 12CA5 抗体的杂交瘤细胞株从学术来源获得，12CA5 抗体抗 HA 表位。

Bradford 染料浓缩液〈!〉

缓冲液 D-

　　0.1 mmol/L EDTA

　　20%甘油（*V/V*）

　　20 mmol/L HEPES-KOH（pH 7.9）〈!〉

　　KCl（0、0.1 mol/L、0.2 mol/L、0.25 mol/L、0.5 mol/L 和 0.85 mol/L）〈!〉

　　临用前，加入

　　　1 mmol/L PMSF〈!〉

　　　提前分批准备含有不同浓度 KCl 的缓冲液 D，临用前加入 PMSF。不要加入 DTT。

庚二亚氨酸二甲酯（DMP；Sigma D8388）〈!〉

洗脱缓冲液

　　0.1 mmol/L EDTA

　　20%甘油（*V/V*）

　　20 mmol/L HEPES-KOH（pH 7.9）〈!〉

　　0.25 mol/L KCl〈!〉

　　临用前，加入

　　　PMSF（0.5 mmol/L 或 1 mmol/L）〈!〉

乙醇胺（0.2 mol/L，pH 7.9）〈!〉

HeLa 细胞核提取物（见方案 12.1）

　　从表达 HA-TBP 的细胞株中制备提取物，该细胞株可从 Berk 实验室获得（Zhou et al. 1992）。

血凝素（HA）肽

　　长期储存用含 0.1 mol/L KCl 的缓冲液 D-稀释到 10 mg/ml。用前用肽稀释缓冲液稀释到 1.5 mg/ml。

KCl（3 mol/L）〈!〉

肽稀释缓冲液

　　0.2 mg/ml 牛血清白蛋白（BSA）

　　0.1 mmol/L EDTA

　　20%甘油（*V/V*）

　　20 mmol/L HEPES-KOH（pH 7.9）〈!〉

　　0.25 mol/L KCl〈!〉

　　0.05% NP-40〈!〉

　　临用前，加入

　　　1 mmol/L PMSF〈!〉

磷酸纤维素（P-11；Whatman，4071-050）

蛋白 A-Sepharose 微珠〈!〉
硼酸钠（0.2 mol/L，pH 9.0）
SDS-PAGE 所需其他试剂

设备

离心机（临床，IEC 或同等物）
层析柱（P-11，30 ml；Whatman，4071-050）
电导计
馏分收集器
微量离心机
针头（小号）
摇床
试管（Falcon，50 ml）
微量离心管（1.5 ml，硅化）
水浴，定在 30℃
SDS-PAGE 附加设备

程序

步骤 1：准备核提取物

核提取物制备，见方案 12.1。从 HA-标记的细胞株制备（Zhou et al. 1992）。

步骤 2：准备 P-11 柱、缓冲液及免疫亲和树脂

准备缓冲液 D-、0.2 mol/L 硼酸钠和 0.2 mol/L 乙醇胺。

准备 P-11 柱

1. 在缓冲液 D-（含 1 mol/L KCl）中准备一个 30 ml 的 P-11 柱。根据制造商的说明对柱进行平衡。

一个 30 ml 的柱子足够用于从 32 L HeLa 细胞制备的 50 ml（约 500 mg）核提取物。可提前 1 天准备柱子并储存在 4℃。

2. 用 120 ml 含 1mol/L KCl 的缓冲液 D-洗柱。

3. 用 60 ml 含 0.5 mol/L KCl 的缓冲液 D-平衡柱。

4. 测定 HeLa 核提取物样品的电导率。加入 3 mol/L KCl，将提取物的电导率调整到与含 0.5 mol/L KCl 的缓冲液 D-的电导率相等。

准备 HA-抗体树脂

该方案使用 0.5 g 微珠，产量大约为 1.0 ml，其规模可作相应调整。

1. 将 0.5 g 蛋白 A-Sepharose 重悬于 5 ml 含有 1.0 mmol/L PMSF 的洗脱缓冲液中，室温下摇动 3 min。

2. 1500*g* 离心 5 min。用 20 ml 含有 1.0 mmol/L PMSF 的洗脱缓冲液洗涤两次。

3. 加入 1～2 ml HA. 11 或 12CA5 抗体（1 mg/ml），室温下摇动 1 h。

4. 1500*g* 离心 5 min，弃上清。

5. 用 10 倍体积（约 20 ml）的 0.2 mol/L 硼酸钠（pH 9.0）洗微珠两次，保存一份 100 μl（约 10 μl 微珠）悬液用于凝胶电泳（“交联前”；见步骤 13）。

6. 3000 r/min 离心 2 min，弃上清。

7. 将微珠重悬于 10 倍体积的 0.2 mol/L 硼酸钠中。

8. 加入 DMP，至终浓度 20 mmol/L（每 20 ml 中加 103.6 mg DMP），室温下摇动 30 min。

9. 用 10 倍体积的 0.2 mol/L 乙醇胺（pH 7.9）洗涤。

10. 用 10 倍体积的 0.2 mol/L 乙醇胺（pH 7.9），室温下孵育 2 h。

11. 用 15 ml 含 1.0 mmol/L PMSF 的洗脱缓冲液洗微珠两次。

12. 将最终洗涤的沉淀重悬于 1 ml 含 0.5 mmol/L PMSF 的洗脱缓冲液，存一份 20 μl样品用于凝胶分析（“交联后”；见步骤 13）。

13. 用 12%的凝胶通过 SDS-PAGE 分析样品，以确证交联。上样 2 μl 交联前的微珠，上样 20 μl 交联后的微珠，监测大、小抗体链（Harlow and Lane 1999）。

步骤 3：纯化表位标记的 TFIID

从核提取物中分级分离 TFIID

1. 将 HeLa 细胞提取物载样到 30 ml P-11 柱上。

2. 用 90 ml 含 0.5 mol/L KCl 的缓冲液 D-洗柱。

3. 用 2 倍柱体积含 0.85 mol/L KCl 的缓冲液 D-缓慢（约 1 ml/min）洗脱柱子，用分部收集器按 3 ml 馏分收集。

4. 用 Bradford 蛋白分析法测定馏分中的蛋白质含量。合并蛋白质浓度最大的馏分（约 15 ml）。注意：仅合并蛋白质浓度最高的馏分。

5. 将洗脱物在含 0.25 mol/L KCl 的缓冲液 D-中透析 2 h。

TFIID 与抗体亲和树脂结合并洗脱 TFIID

1. 每毫升透析物加入 100 μl 与抗体交联的蛋白 A-Sepharose 微珠。注意：使用硅化的 1.5 ml 离心管，以避免蛋白质丢失。

2. 在振荡器上 4℃孵育 4 h 至过夜，使蛋白质与抗体偶联微珠结合。

3. 在微量离心机中全速（14 000 r/min）离心 15 s，以沉淀微珠。

4. 用含 0.25 mol/L KCl 的缓冲液 D-洗微珠 3 次。

a. 吸取 1 ml 缓冲液加到蛋白质/微珠浆液上，轻柔涡旋。

b. 在冰上孵育 5 min。

c. 在微量离心机中以最大转速离心微珠 15 s，弃上清。

5. 将 HA 肽混合液（1.5 mg/ml 溶于肽稀释缓冲液）与微珠按 1∶1 比例混合（即 125 μl 微珠洗脱物与 125 μl HA 肽混合液混合）。

6. 室温（约 25℃）下摇动 20 min，孵育 HA/微珠混合液。

7. 30℃水浴中孵育 HA/微珠混合物 20 min。

8. 用乙醇擦拭试管外壁，用小规格注射针在管的底部扎一个非常小的孔。

目的是使液体流出，而微珠保留下来，因此只能用针刚刚扎透试管。

9. 将含有微珠的 1.5 ml 微量离心管置于另一个 1.5 ml 管中，用胶带或蜡膜将两管固定在一起。

10. 将两个管置于 50 ml Falcon 管中。用临床离心机 4℃离心 4 min，或直到微珠干燥。

11. 将洗脱的 TFIID 分装成 20 μl 的样品，此阶段可以加入 DTT 至 1 mmol/L。

步骤 4：通过 SDS-PAGE 和免疫印迹法分析纯化的 TFIID

1. 滴定被洗脱的 TFIID。与起滴定作用的、纯化的分子质量标准物蛋白质一起在 10％SDS 凝胶上分级分离。

   也可用银染的 SDS 凝胶确证 TFIID 的亚基构成。

2. 通过 Western 印迹法分析蛋白质。在蛋白质检测中，利用抗 HA 标签的抗体 HA.11（1∶500 稀释）作为初级抗体。

   其他免疫印迹可用抗 TBP 和不同 TAF 的抗体（Santa Cruz Biotechnology，Inc）进行。

附加说明

1. *与免疫亲和柱的结合*。重要的是在 P-11 洗脱物与免疫亲和树脂结合前后监测提取物。至少 1/2 或 2/3 的 TFIID 应该结合。这一点在比较开始提取物和缺失提取物时很明显。同样，当从免疫亲和树脂中洗脱时，在显示的条件下预计只有 1/2～2/3 的 TFIID 被洗脱。较高的盐浓度（即 0.5～1.0 mol/L）有利于更定量的洗脱，但不与 DNA 结合实验直接兼容。

2. *TFIID 浓度*。一般情况下，DNase I 足迹或体外转录实验中使用 200 ng 左右的免疫纯 TFIID，这相当于 1～3 μl 亲和柱洗脱物。300 mg 核提取物的一般产量最多为 20 μg，通常更少。如果 TFIID 以更稀的浓度洗脱，虽然这种蛋白质仍可用于体外转录和 EMSA，但不可能得到清晰的 DNase I 足迹。通常，可以通过从 *E. coli* 中制备重组 TBP 来确定 TFIID 的摩尔浓度。通过以 TBP 抗体（Santa Cruz Biotechnology 公司）进行的免疫印迹，将已知摩尔量的 TBP 和 TFIID 比较。当尝试保存 TFIID 时，我们采用这种方法估算 TFIID 浓度，而不采用微量 Bradford 分析或银染凝胶。

3. *组分*。偶尔使用银染 SDS 凝胶确证 TFIID 的亚基组成。此外，也可以用从 Santa Cruz Biotechnology 公司获得的抗 TBP 和多种 TAF 的抗体进行免疫印迹分析。

4. *蛋白 A-微珠污染*。从 TFIID 制剂中去除所有的蛋白 A-微珠是极其重要的。微珠污染会在利用 TFIID 的足迹和 EMSA 反应中造成问题。一般情况下，轻弹管壁可以看到微珠，在这种情况下，在微量离心机离心纯 TFIID，并小心转移上清至另一个管中，但须避免吸上微珠沉淀物。

疑难解答

TFIID 洗脱浓度太低，无法用于生化分析

可能原因：低浓度的起始材料或低效价的抗体是该程序失败的两个最常见原因。

解决办法：确保 P-11 TFIID 峰材料尽可能浓缩。尝试只结合来自 P-11 柱的最浓组分。测试 HA. 11 或 12CA5 抗体的效价，并将结果与商业来源的样品或由其他研究人员提供的样品比较。

TFIID 从树脂上未洗脱

可能原因：这是一个常见问题，通常与变质的肽有关，或在浓度太低时出现，或在洗脱中使用低盐浓度。

解决办法：尝试新肽和稍高（增加 100 mmol/L）的盐浓度。

方案 12.5　从表达 FLAG 标记调解因子亚基的 HeLa 细胞系中纯化调解因子

调解因子（Med）复合物在由 RNA Pol II 转录的启动子特异性激活中发挥关键作用。我们之所以提供 Med 的纯化方案，是因为它是激活因子的一个主要靶标，并且它可以用于许多类型的亲和结合和固定化模板研究，以评估单个激活因子之间的相互作用。此外，所有的 Med 亚基都已经被克隆并可用于单个的结构-功能分析，以了解一个特定的激活因子如何与特定亚基相互作用。在酵母中的研究将 Med 初步鉴定为活化转录所必需的因子（Kelleher et al. 1990；Flanagan et al. 1991）。酵母 Med 在体外和体内也刺激基础转录（Kim et al. 1994）。随后的研究使用传统的层析技术搜寻了哺乳动物的 Med 复合物。这些哺乳动物制剂基于它们与共激活因子的缔合、它们与酵母 Med 相关的组成，或它们对转录的正效应（Jiang et al. 1998；Sun et al. 1998；Gu et al. 1999；Ito et al. 1999；Naar et al. 1999；Rachez et al. 1999；Ryu et al. 1999）。最初，基于报道的亚基组成，认为 Med 在组成上变化多端，后来的研究使用 FLAG 标记 Med 亚基产生了更高纯度的 Med 复合物（Sato et al. 2004）。标签不同的亚基导致了相似组成的 Med 制剂，尽管似乎有两个复合物，它们含有 CDK8、CDK11 或 Pol II，但不清楚它们是否相互排斥。表 12.1 罗列了加标签 Med 亚基的细胞系。Roeder 实验室已生成 FLAG 标记 CDK8 和 Nut2（Gu et al. 1999；Malik and Roeder 2000）。Conaway 实验室已生成其他几个 FLAG-标记亚基（Sato et al. 2004）。该方案提供了一个从稳定表达 FLAG 标记兼性蛋白（Intersex）的 HeLa 细胞中生产纯化的 Med 复合物的方法。

时线与安排

第 1 天：从稳定表达 FLAG 标记 Med 亚基的细胞系中制备核提取物。

第 2 天：准备缓冲液和实验室器具，4℃预冷。

第 3 天：纯化 FLAG 标记间性蛋白亚基。

提纲

步骤 1：从表达 FLAG 标记兼性蛋白的 HeLa 细胞系中制备核提取物（见方案 12.1）（使用低浓度的 0.1 mmol/L DTT，或不用 DTT）。

步骤 2：准备缓冲液并在 4℃预冷（1 h）。

步骤 3：4℃平衡抗 FLAG 树脂（0.5 h）。

步骤 4：4℃将 FLAG-Med 结合到抗 FLAG 树脂上并洗涤（6 h）。

步骤 5：4℃从树脂洗脱 FLAG-Med（1.5 h）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

通常从缓冲液中省略 DTT，因为它会从微珠上解离 FLAG 抗体的二硫键。用 0.2 μm的瓶口过滤膜过滤缓冲液。

琼脂糖微珠（抗-FLAG-M2-偶联；Sigma，A2220）

琼脂糖微珠以 1～2 倍工作体积的小等份（200～400 μl，配成含 50%甘油的 50%微珠浆液）储存在－20℃，也可有 4℃储存的微珠。

抗 FLAG 抗体（Sigma，F1804）（可选）

缓冲液 A

10 mmol/L HEPES（pH 7.9）

10 mmol/L KCl〈!〉

1.5 mmol/L $MgCl_2$〈!〉

临用前，加入

0.5 mmol/L PMSF〈!〉

蛋白酶抑制剂（1 片；Roche，11836170001）

干冰（或液氮）〈!〉

洗脱缓冲液

10 mmol/L HEPES（pH 7.9）

300 mmol/L KCl〈!〉

1.5 mmol/L $MgCl_2$〈!〉

0.05% NP-40〈!〉

临用前，加入

0.5 mmol/L PMSF〈!〉

蛋白酶抑制剂（1 片；Roche，11836170001）

300 μg/ml FLAG 肽（Sigma，F3290）。

储存小份 FLAG 肽于－80℃，避免反复冻融，过滤缓冲液后加入肽。

HeLa 细胞核提取物（见方案 12.1）

从表达 FLAG 标记兼性蛋白（或其他 FLAG 标记亚基；Sato et al. 2003）的细胞系中准备核提取物。在 FLAG 纯化之前不需要透析提取物，除非 HeLa 提取物中使用了 DTT。如果是这样，应使用无 DTT 的缓冲液 D-（0.1 mol/L KCl）透析（见方案 12.4）。

洗涤缓冲液

10 mmol/L HEPES（pH 7.9）

300 mmol/L KCl〈!〉

1.5 mmol/L $MgCl_2$〈!〉

0.05% NP-40〈!〉

临用前，加入

0.5 mmol/L PMSF〈!〉

蛋白酶抑制剂（1 片；Roche，11836170001）

SDS-PAGE 和（或）免疫印迹所需的其他试剂。

设备

离心机（临床；IEC 或同等物）

离心机（Beckman 或 Sorvall）

过滤器（0.2 μm 瓶顶过滤器；Nalgene，291-4520）

转头

试管（锥形，15 ml）

微量离心管（硅化，1.5 ml）

SDS-PAGE 和（或）免疫印迹附加设备

程序

步骤 1：准备核提取物

关于核提取物制备，见方案 12.1。从表达 FLAG 标记兼性蛋白的 HeLa 细胞系中准备核提取物（Sato et al. 2003）。

步骤 2：准备缓冲液并在 4℃ 冷却

用前在 4℃预冷缓冲液、硅化试管、锥形管及吸头。准备缓冲液 A、洗涤缓冲液和洗脱缓冲液。注意：DTT 往往从缓冲液中省略，因为 DTT 从微珠上解离 FLAG 抗体的二硫键。

步骤 3：4℃ 平衡抗 FLAG 树脂

1. 用冰冷的洗涤缓冲液（每 6 ml 核提取物准备 200 μl 50%的微珠浆液）4℃下平衡抗 FLAG 抗体偶联的琼脂糖微珠 30 min。

微珠用量可能要依提取物中 FLAG-Med 的表达水平做经验性地改变。

2. 将微珠浆液 1000 *g* 离心 5 min，除去上清。

3. 用 1ml 洗涤缓冲液重悬微珠，1000 *g* 再次离心 5 min，除去上清。

4. 重复步骤 3：3 两次。

5. 用洗涤缓冲液重悬微珠成 50%的浆液。

200 μl 50%的微珠浆液应产生 100 μl 压实珠体积（PBV）。

步骤 4：4℃下 FLAG-Med 与抗 FLAG 树脂的结合和洗涤

1. 解冻约 6 ml 从表达 FLAG-兼性蛋白的 HeLa 细胞系准备的未透析的核提取物。4℃，20 000 *g* 离心 20min，清除沉淀物质或不溶性碎片。

2. 将含有核提取物的上清液转移到 15 ml 锥形管，置于冰上。

3. 加入 0.4 倍核提取物体积（即 6 ml 核提取物加 2.4 ml）的冰冷的缓冲液 A，调整核提取物盐浓度到 0.3 mol/L KCl。

4. 每 6 ml 核提取物加 200 μl 平衡过的抗 FLAG-琼脂糖微珠浆液（来自步骤 3：5），4℃旋转孵育 4 h。

5. 4℃，1000 *g* 离心 5 min，取出并保存上清（即未结合的“穿透物”），用于凝胶分析，以确定 FLAG-微珠结合相对于输入的效率。要获得 Med 耗尽的提取物，需进行至少三次抗 FLAG 免疫沉淀。为了优化通过树脂的次数，须用抗 FLAG 免疫印迹监测耗竭过程。

6. 微珠用 50×PBV 体积的洗涤缓冲液（如每 100 μl PBV 加 5 ml 洗涤缓冲液），4℃旋转孵育 5 min。

7. 4℃，1000 *g* 离心 5 min，移除上清液。

8. 重复步骤 4：6 和 4：7 共 4 次（共洗涤 5 次）。

步骤 5：4℃下从树脂上洗脱 FLAG-Med

1. 用 5×PBV 体积洗涤缓冲液（100 μl PBV 使用 500 μl 洗涤缓冲液）重悬微珠，转移到 1.5 ml 硅化管中。

2. 用另外的 5×PBV 体积洗涤缓冲液清洗 15 ml 锥形管，收集任何剩余的微珠，将其加入到 1.5 ml 硅化管中。1000 *g* 离心 2 min，弃去上清液。

3. 每 100 μl 微珠加入 100 μl 洗脱缓冲液，4℃旋转孵育 30 min。

4. 4℃，1000 *g* 离心 2 min，保留含有 Med 复合物的上清。

5. 重复步骤 5：3 和 5：4 两次（共三次）。

6. 为了确定纯度，在聚丙烯酰胺凝胶上上样洗脱产物的滴定用于银染。
   洗脱产物也可用 FLAG 抗体进行免疫印迹分析，兼性蛋白的大小约为 26 kDa。

7. 将洗脱的 FLAG-Med 分装成 20～50 μl 样品，干冰或液氮冻结。

附加说明

1. 洗脱。有必要分析洗脱，以确定纯度。通过向用于银染的聚丙烯酰胺凝胶上载样洗脱产物的滴定可以很轻松地做到这一点。此外，还建议用抗 FLAG 抗体免疫印迹法分析洗脱产物。兼性蛋白分子质量大小约为 26 kDa。

2. 抗 FLAG 琼脂糖微珠和肽。微珠应存放成 1 或 2 个工作体积（此方案使用 200～400 μl 50%的微珠浆液）的小等份，保存在－20℃。也可购买 4℃储存的微珠。此外，FLAG 肽应分装成小等份储存于－80℃。

3. Med 从提取物中耗尽。为了从 HeLa 核提取物中耗竭 Med 复合物，有必要使提取物通过抗 FLAG 微珠至少三次。有必要优化通过抗 FLAG 微珠的次数，并用抗 FLAG 免疫印迹监测耗竭。

疑难解答

与抗 FLAG 微珠没有结合

可能原因：缓冲液中可能有过量的 DTT。

解决方案：制备新鲜的小体积微珠等份。确保在任何缓冲液中 DTT 的浓度小于 0.1 mmol/L，因为这将破坏微珠偶联抗体的二硫键。

Med 不能有效地从抗 FLAG 微珠上洗脱

可能原因：FLAG 肽已经冻融次数太多或洗脱条件非最佳。

解决方案：制备新鲜的 FLAG 肽等份。也可以尝试用 3×FLAG 肽（Sigma，F4799）洗脱。这通常是用于表达 3×FLAG 肽的蛋白质，但它可以增加表达一个 FLAG 肽蛋白的洗脱效率。也可以尝试增加孵育时间或提高 FLAG 肽浓度到 500 μg/ml。洗脱也可在室温下和较高的盐浓度下进行，以提高产量。

方案 12.6　固定化模板分析

启动子 DNA 上大约 4 MDa 的 RNA Pol II PIC 的组装对转录激活至关重要。形成增强体的多个激活因子之间、激活因子与 PIC 之间，以及 PIC 内蛋白质之间的蛋白质-蛋白质相互作用都稳定最终的复合物（Carey 1998）。PIC 组装包括但不限于以下一些元件：① GTF，包括 TFIID 的 TBP、TFIIA、TFIIB、TFIIE、TFIIH 和 TFIIF（Woychik and Hampsey 2002；Sims et al. 2004；Roeder 2005）；②共激活因子，包括 30 亚基 Med 复合物（Boube et al. 2002）和 TFIID 的 TAF（Hahn 1998；Chen and Hampsey 2002）；③各种染色质重塑因子（Peterson and Laniel 2004；Smith and Peterson 2005；Verdone et al. 2005）；④RNA Pol II（Woychik and Hampsey 2002）。研究相对小蛋白质的 DNA 协同结合问题的传统方法包括 DNase I 足迹和 EMSA 分析。Sharp 及其同事的工作为使用这些技术研究激活因子存在情况下通用转录因子的逐步组装提供了范例（Buratowski et al. 1989）。虽然 EMSA 和 DNase I 足迹具有明显的优势，但对于使用大型复合物而言，这些技术都是有限的。包含多个大型激活因子（如 MED 和 TFIID）的复合物，因太大而无法用 EMSA 解析。

为了绕过这个限制，固定化模板募集分析已经被有效地用于研究这些大型多蛋白复合物的协同结合（Johnson et al. 2004）。固定化模板已被成功地用于在确定的体外系统中观察蛋白质相互作用。Kingston 研究小组已采用固定化染色质阵列来表征 GAGA 和 Zeste 蛋白对多梳蛋白家族复合物在同源异型基因上的抑制功能的影响（Mulholland et al. 2003），以及观察多梳蛋白组复合物的抑制机制（Lavigne et al. 2004）。Hahn 研究小组采用固定化

模板观察 Pol II-TFIIB 在 PIC 上的定位（Chen and Hahn 2004）。Roeder 研究小组使用固定化模板分析得出了这样的结论，即转录起始后，激活因子招募的 PIC 显示了 Med 的部分瓦解，而这种瓦解导致了类 PC2 复合物被保留在支架中（Malik et al. 2005）。

在固定化模板分析中，生物素化的 DNA 片段是通过连接到链亲和素包被的磁性 Dynal™微珠上而被固定的。DNA 结合蛋白与固定化模板孵育，然后，通过用磁性颗粒集中器分离固定化模板，以及通过随后的“磁珠/模板”清洗，去除没有特异性组装在复合物中的蛋白质。结合的蛋白质用免疫印迹检测。为了检测协同结合，两种或两种以上蛋白质或蛋白质复合物以亚饱和量加入固定化模板。一种蛋白质和协同结合伙伴与模板的共同结合，可与每个单个蛋白质或蛋白质复合物单独的结合相比较。

用固定化模板分析来研究蛋白质复合物的结合提供了传统方法不具有的很多优势。固定化启动子模板也可以用于功能分析，如体外转录，从而使复杂的装配与功能建立联系。此外，由于蛋白质结合是通过免疫印迹分析的，可以检测一个多亚基复合物中单个亚基的存在，不论多亚基复合物有多大。大型复合物组装的分析可以在很短的时间内实现，并且不再牵扯精密设备。因为蛋白质检测方法是免疫印迹，依靠电泳迁移率变动检测蛋白质复合物的个别亚基的翻译后修饰也是可能的，而且由于免疫印迹相对敏感，通过使用标准曲线，可重复性检测 DNA 结合中微小（小于两倍）的改变。转录复合物组装研究的特别点是，固定化 DNA 模板可与核小体预组装，以模拟细胞内的更自然的染色质环境（Hassan et al. 2001）。

固定化模板分析的缺点如下。蛋白质结合的检测依赖于是否可获得高滴度抗体。此外，背景非特异性 DNA 结合，如蛋白质结合 DNA 末端，正像 RNA Pol II 的特点，会带来问题。由于非特异性 DNA 结合是这一技术的共性问题，在实验设计中应包括特异性对照。例如，无 TATA 的启动子的使用应作为 TFIID 特异结合的对照。此外，和 DNase I 足迹相比，固定化模板分析提供了较少的位置信息，并且关于功能与募集的相关性，组装的 PIC 复合物的活性远低于 100%，其中的原因我们并不完全了解。

时线与安排

在使用固定模板系统前，必须有蛋白质和（或）纯化的核提取物。生成固定化模板很简单，因为它只需要 PCR 扩增并将 PCR 扩增产物结合到磁珠上。这个过程能在短短的一天内完成。但是，一旦纯化的蛋白质和核提取物被加入系统，结合反应就会变得复杂。必须单独滴定每种蛋白质以了解加入每个反应的饱和和亚饱和量。此外，固定化模板上的转录反应效率会大大地降低，因此转录分析应首先在超螺旋质粒 DNA 上优化。当进行此分析时，我们通常会使用清楚表征的模式激活因子，如 Gal4-VP16。因此，当我们在下文提到激活因子时，一般是指 200 ng/μl 的 Gal4-VP16 原液。

固定化模板分析

第 1 天：生成生物素化的 DNA 片段并连接到磁珠上。

第 2 天：蛋白质结合反应或转录分析。

第 3 天：免疫印迹和（或）数据分析。

提纲

步骤 1：准备工作：设计引物。

步骤 2：制备 PCR 片段（3 h）。

- PCR 扩增。
- 电泳和凝胶纯化。

步骤 3：将 DNA 片段固定到磁珠上（2 h）。

- DNA 结合到 Dynal 磁珠。
- 磁珠结合 DNA 的定量。

步骤 4：实验：蛋白质结合分析（8 h）。

- 蛋白质结合反应。
- 激活因子从核提取物募集共激活因子蛋白。
- 激活因子募集纯化的共激活因子蛋白。

步骤 5：实验：转录分析（12 h）。

- 转录反应。
- 引物延伸。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

激活因子蛋白（200 ng/μl，如 Gal4-VP16）

抗体

　　用于免疫印迹检测激活因子蛋白和共激活因子的抗体可商业化获得（例如，Santa Cruz Biotechnology，Upstate Group，Inc.）

牛血清白蛋白（BSA；50 μg/μl）

2× B/W 缓冲液

　　1 mmol/L EDTA

　　2 mol/L NaCl

　　10 mmol/L Tris-HCl（pH 7.5）

缓冲液 D

　　0.1 mol/L EDTA

　　20% 甘油

　　20 mmol/L HEPES（pH 7.9）

　　0.1 mol/L KCl〈!〉

　　临用前，加入

　　　　1 mmol/L DTT〈!〉

　　　　1 mmol/L PMSF〈!〉

DNA 梯状条带（1 kb，New England BioLabs，323-2L）

　　梯状条带应包含与从磁珠上消化的 DNA 大小相同的 DNA 条带。

dNTP 混合物（25 mmol/L）

磁珠 Dynabeads ® M-280 链亲和素（Invitrogen，112-05D）

　　这些磁珠最适合结合小于 1 kb 的 DNA 片段。要结合长度超过 1 kb 的 DNA 片段时，使用 Dynal kilobase BINDER 试剂盒。

凝胶提取试剂盒（如 QIAquick；Qiagen，28704）

HeLa 细胞核提取物（见方案 12.1）

Med 蛋白（纯化的，300 ng/μl，见方案 12.5）

$MgCl_2$（0.1 mol/L）〈!〉

$MgCl_2$（25 mmol/L）（可选）〈!〉

矿物油（可选）

NP-40（1%）〈!〉

NTP 混合物（0.5 mmol/L）

pGEM3（100 ng/μl）

质粒 DNA（20 ng/μl）

引物（5′端和 3′端；10 pmol/μl）

限制酶

10×限制酶缓冲液

Taq 酶（5 U/μl）

10× *Taq* 反应缓冲液

TE 缓冲液

洗涤缓冲液

　　0.625×缓冲液 D

　　7.5 mmol/L $MgCl_2$〈!〉

　　0.05% NP-40〈!〉

　　临用前，加入

　　　1 mmol/L DTT〈!〉

　　　1 mmol/L PMSF〈!〉

琼脂糖凝胶电泳其他试剂；SDS-PAGE、免疫印迹试剂和转录分析试剂（见方案 12.2）

设备

离心机（台式）

磁性颗粒集中器（MPC；Invitrogen，120-20D）

PCR 扩增热循环仪

移液器和吸头

转头

微量离心管（1.5 ml）

PCR 管

紫外灯（低功率）〈!〉

进行琼脂糖凝胶电泳、SDS-PAGE 和免疫印迹，以及转录检测（见方案 12.2）所需的附加设备。

程序

步骤 1：准备工作：设计引物

设计引物，使得每个引物约为 25 bp。合成用生物素标记的 5′引物并使其靠近 5′端含有特定限制酶酶切位点。为获得最佳的生物素化 PCR 扩增片段与链亲和素包被磁珠的结合，PCR 产物应小于 1 kb。此外，激活因子结合位点应该距 5′端约 200 bp，以便让大的蛋白质复合物结合 DNA 而没有来自珠子的空间位阻的影响。

步骤 2：制备 PCR 片段

PCR 扩增

通常情况下，每两个微珠结合反应我们进行下述 4 个独立 PCR。

1. 在 PCR 管中，准备如下反应：

10×*Taq* 反应缓冲液	10 μl
5′引物	1 μl
3′引物	1 μl
dH_2O	85 μl
质粒 DNA	1 μl
dNTP 混合物	1 μl
Taq 酶	1 μl
总体积	100 μl

如果反应缓冲液中没有 $MgCl_2$，加入 1～2 μl 25mmol/L $MgCl_2$。

2. 用矿物油覆盖反应混合物（除非热循环仪有热盖）。
3. 在热循环仪中执行下面的扩增程序：将退火温度设定得低于引物熔解温度。

步骤	温度/℃	时间/min	
1	95	5：00	
2	95	0：45	30 个循环
3	45	0：45	30 个循环
4	72	1：00	30 个循环
5	72	5：00	
6	4	持续	

电泳和凝胶纯化

1. 用 1.5%琼脂糖凝胶在 TBE 中电泳 PCR 混合物，溴化乙锭染色。
2. 在紫外线下观察条带，从琼脂糖凝胶切出适当大小的条带。

3. 按凝胶回收试剂盒说明书所述的方法纯化 DNA，将纯化的片段洗脱在 30 μl 试剂盒提供的缓冲液 EB 中。

步骤 3：固定 DNA 片段

DNA 与 Dynal 磁珠结合

在片段 PCR 扩增不可能或不佳的情况下，用生物素化核苷酸末端标记限制性片段，并使用它替代 PCR 片段。

1. 在两个 1.5 ml 微量离心管各加 30 μl 磁珠浆液，将离心管置于磁性颗粒集中器（MPC）中。

   集中器将磁珠吸附到离心管壁。

2. 经过约 30 s，用移液器从管底移除缓冲液。
3. 用 200 μl 2×B/W 缓冲液洗磁珠两次，之后将磁珠重悬于 60 μl 2×B/W 缓冲液中。
4. 向两组磁珠各加入两个 PCR 反应扩增的重悬于缓冲液 EB 的 DNA 片段（来自步骤 2：3）。混合物在旋转混合器上室温下孵育 1 h。
5. 用 MPC 去除上清。磁珠先用 200 μl 1×B/W 缓冲液洗两次，再用缓冲液 D 洗两次，最后重悬于 60 μl 缓冲液 D 中。

   磁珠可在 4℃下储存长达 2 个月。磁珠可用于蛋白质结合分析、使用核提取物或纯化蛋白进行的募集分析，以及转录分析。

使用唯一 5′端限制性酶切位点定量磁珠结合的 DNA

1. 准备限制酶混合物：

限制酶	1 μl
10×限制酶缓冲液	2 μl
dH_2O	17 μl
总体积	20 μl

2. 用 TE 洗涤 5 μl 磁珠结合的 DNA 两次，并将其重悬于限制酶混合物中。
3. 根据限制酶提供的说明进行孵育。为确保完全切割，定期混合反应，防止磁珠沉淀。
4. 磁分离并转移出上清液。将 2 μl、4 μl 和 8 μl 限制酶消化液稀释在电泳样品缓冲液中。

   这大约相当于 0.5 μl、1.0 μl 和 2.0 μl 与磁珠结合的 DNA。

5. 在 1.5%琼脂糖凝胶上电泳样品。为了定量分析，在凝胶上载样 2.5 μg、5.0 μg 和 10.0 μg DNA 梯状条带。650 bp 片段的一般产量是每微升磁珠 20～50 fmol DNA。

步骤 4：蛋白质结合分析

蛋白质结合反应

蛋白结合反应可包括纯化的因子和（或）核提取物。第一个实验应证实已知与固定化片段中的 DNA 序列结合的激活因子的结合。成批准备反应混合物，在最后加入激活因子蛋白前分装。

1. 在冰上，准备如下一批足够用于 6 个反应的混合液：

缓冲液 D	219 μl
dH_2O	84 μl
0.1 mol/L $MgCl_2$	27 μl
1% NP-40	91 μl
BSA	6 μl
pGEM3	12 μl
总体积	354 μl

pGEM3 作为竞争物，以防止蛋白质与 DNA 的非特异性相互作用。

2. 准备激活因子的三倍稀释系列：4 μl 缓冲液 D 中加入 2 μl 未稀释激活因子原液（200 ng/μl），以获得“激活因子”、“1∶3 激活因子”、“1∶9 激活因子”和“1∶27 激活因子”。

3. 准备 5 个微量离心管，每个加 1 μl DNA 模板结合的磁珠（约 40 fmol），放在 MPC 上去除上清。立即加入成批准备的混合液，小心不要使磁珠变干。

4. 完成如下反应混合液。最后加入激活因子蛋白：

反应	1	2	3	4	5
磁珠（成批准备的混合液）	59 μl	59 μl	59 μl	59 μl	59 μl
缓冲液 D	1 μl	—	—	—	—
激活因子	—	1μl	—	—	—
1∶3 激活因子	—	—	1 μl	—	—
1∶9 激活因子	—	—	—	1 μl	—
1∶27 激活因子	—	—	—	—	1 μl
总体积	60 μl	60 μl	60 μl	60 μl	60 μl

5. 轻轻混合，重悬磁珠，在室温下旋转孵育 30 min。

6. 用 200 μl 洗涤缓冲液洗涤 3 次。

7. 用 10 μl SDS-PAGE 载样染料重悬磁珠。9%丙烯酰胺 SDS-PAGE 电泳。激活因子蛋白免疫印迹分析。

来自核提取物的共激活因子蛋白的激活因子募集

在核提取物存在的情况下，激活因子蛋白被结合到模板上，经过一系列洗涤去除未结合的蛋白质。结合蛋白用免疫印迹分析。下面的实验是一个基本的设置，可能需要激活因子蛋白或核提取物的进一步滴定。

1. 准备如下一批足够用于 7 个反应的混合液：

缓冲液 D	143.5 μl
dH_2O	94.5 μl
0.1 mol/L $MgCl_2$	31.5 μl
1% NP-40	10.5 μl
BSA	7.0 μl
pGEM3	14.0 μl
总体积	301.0 μl

2. 稀释激活因子蛋白，获得最佳结合浓度（如步骤 4：1～4：7 那样确定）。

3. 使核提取物达到室温（约 1 min）。台式离心机最高转速离心 1 min，以使沉淀的蛋白质结块，这可以减少蛋白质的非特异性结合。

4. 准备 6 个微量离心管，每个加 1 μl DNA 模板结合磁珠（约 40 fmol），放在 MPC 上去除上清，立即加入成批准备的混合液，小心不要使磁珠变干。

5. 完成如下反应混合液：

反应	1	2	3	4	5	6
磁珠磁珠（成批准备的混合液）	43 μl	43 μl	43 μl	43 μl	43 μl	43 μl
缓冲液 D	17 μl	16 μl	12 μl	8 μl	—	1μl
激活因子（如 1：27）	—	1 μl	1 μl	1 μl	1μl	—
核提取物	—	—	4 μl	8 μl	16 μl	16 μl
总体积	60 μl	60 μl	60 μl	60 μl	60 μl	60 μl

缺乏激活因子或核提取物的对照对于正确解释提取物中因子的背景结合和确保提取物没有移除激活因子是必不可少的。

6. 轻轻混合，重悬磁珠。在室温下旋转孵育 30 min。

7. 用 200 μl 洗涤缓冲液洗涤 3 次。

8. 用 10 μl SDS-PAGE 载样染料重悬磁珠。9%丙烯酰胺 SDS-PAGE 电泳。取结合实验中所用核提取物量的 1%～10%作为蛋白质样品直接点在胶上，以确定各种共激活因子或 GTF 被募集的效率。

9. 免疫印迹分析激活因子蛋白和共激活因子蛋白（如 Med 亚基、GTF、CBP/p300 蛋白，或 Swi/Snf 亚基）。

纯化的共激活因子蛋白的激活因子募集

依靠激活因子，使用固定化模板可研究激活因子与特定共激活因子蛋白的相互作用。这里描述的程序是一个基本方案，可能需要激活因子蛋白、Med 或 p300 的进一步滴定。

1. 按步骤 4：1～4：6 所述进行激活因子蛋白与 DNA-结合磁珠的预结合。按步骤 4：4 的确定，使用最佳用量的激活因子。作为对照，同步准备一套不加激活因子的磁珠。

2. 准备如下一批足够用于 10 个反应的混合液：

试剂	用量
缓冲液 D	290 μl
dH_2O	120 μl
0.1 mol/L $MgCl_2$	45 μl
1% NP-40	15 μl
BSA	10 μl
pGEM	20 μl
总体积	500 μl

3. 准备 8 个离心管，4 个加 1 μl 的激活因子结合的磁珠，4 个加对照磁珠，放在 MPC 上去除上清。立即加入成批准备的混合液，小心不要使磁珠变干。

4. 完成如下反应混合液：

反应	1	2	3	4
磁珠（成批准备的含有或不含有激活因子蛋白的混合液）	50 μl	50 μl	50 μl	50 μl
缓冲液 D	10 μl	9 μl	7 μl	1 μl
纯化的 Med	—	1 μl	3 μl	9 μl
总体积	60 μl	60 μl	60 μl	60 μl

5. 轻轻混合，重悬磁珠。在室温下旋转孵育 30 min。
6. 用 200 μl 洗涤缓冲液洗涤 3 次。
7. 用 10 μl SDS-PAGE 载样染料重悬磁珠。9%丙烯酰胺 SDS-PAGE 电泳。
8. 免疫印迹分析激活因子蛋白和 Med 亚基。

步骤 5：转录分析

转录反应

与超螺旋质粒 DNA 相比，固定化模板的转录效率约减少 5 倍。

1. 建立转录反应，程序同固定化模板蛋白结合分析（步骤 4：1 和步骤 4：2）。该批混合液中加入 0.5 mmol/L 每种 NTP。

2. 按方案 12.2 所述进行转录反应。

   使用 MPC 沉淀磁珠，去除上清；此处不能用离心的方法沉淀磁珠。注意终止缓冲液加蛋白酶 K 处理后，只有上清液需要执行其余的步骤。

引物延伸

按方案 12.2 所述进行引物延伸反应。

附加说明

1. 上面使用的 Dynal 磁珠最适合结合小于 1 kb 长的 DNA 片段。Invitrogen 提供的“Dynal kilobaseBINDER 试剂盒”用于结合超过 1 kb 的 DNA 片段。

2. 通过如下措施可优化结合反应条件。
   - 增加盐浓度，创造更严谨的结合条件。
   - 增加洗涤步骤的次数，减少观察到较弱的相互作用。
   - 改变反应孵育时间，观察不同蛋白质的最大结合。
   - 增加竞争 DNA 用量，减少非特异性结合。
   - 使用线性竞争 DNA，减少蛋白质对固定 DNA 自由端的非特异性结合。
   - 增加 BSA 浓度，防止非特异性蛋白质相互作用。
   - 减少体积，促进低浓度蛋白质的结合。

3. 固定化模板是研究协同相互作用的有力工具。要开始研究这些属性，重要的是要确定亚饱和蛋白结合条件，这可以通过在模板上滴定单个蛋白质和改变结合条件的严谨性来实现。然后，用每种单个蛋白质和全部蛋白质建立反应。为了测量协同性的刺激效应，在进行免疫印迹时，必须包括蛋白质的输入滴定。

4. 记住，蛋白质可以与模板结合，紧接着洗去未结合的蛋白质，然后第二个蛋白质或提取物与模板结合。蛋白质预结合可以让模板饱和，同时消除转录压制。

5. 可以很容易地将固定化模板用于研究结合动力学。可通过将反应放置在 MPC 上、洗涤并将反应重悬于载样染料而快速终止时间进程反应。一般的时间进程实验可能包括 1 min、3 min、10 min、30 min、90 min 反应。某些蛋白质的最大结合可能在 3 min以内（如 GAL4-VP16），而对于某些蛋白质（如 TFIID）可能依实验条件的不同需要长达 30 min。

6. 对 Western 印迹募集分析而言，通过蛋白质转移后印迹膜的切割，一个结合反应可用于免疫印迹几种蛋白质。要充分利用这一优势，首先应确定用于免疫印迹的蛋白质的分子质量，其次应至少将两个泳道载样蛋白质梯状条带作为分子质量参考点，以便直线切膜。

疑难解答

高背景蛋白质结合

可能原因：非最佳结合条件。

解决方案：通过改变单个结合条件优化。要建立更严紧的结合条件，尝试增加盐或去污剂浓度，或改变孵育温度。为了减少非特异性结合相互作用，尝试增加竞争 DNA 的量，采用线性竞争 DNA，或减少孵育时间。此外，加入固定模板前，在竞争 DNA 和 BSA 存在的条件下将反应混合物孵育几分钟，可以降低背景结合。另一种可利用的方法是滴定每个反应中使用的固定化模板的量。

目的蛋白的低结合

可能原因：非最佳结合条件和免疫印迹造成的灵敏度缺乏。

解决方案：这可能需要纯化目的蛋白。如果蛋白质已被纯化，尝试使用新制备的蛋白质。优化实验，创造低严紧性结合条件。这可能是抗体的灵敏性不足以检测低水平蛋白质，或 Western 印迹条件与特定抗体不相容。尝试使用不同的封闭剂，如牛奶或 BSA。

弱转录

可能原因：链亲和素包被的磁珠抑制转录。

解决方案：用启动子区制备一新的模板，使启动子区进一步远离磁珠。此外，如果磁珠连在模板的 5′端，而不是 3′端上，转录产量会更高。

方案 12.7 Pol II 开放复合物的高锰酸钾探测

衡量激活因子刺激起始前复合物组装的能力的方法之一，是检测其对 Pol II 开放复合物的刺激，显示开放复合物的常用试剂是高锰酸钾（$KMnO_4$）。该方法是由 Gralla 及其同事于 1989 年创立的，当时用于研究原核 RNA 聚合酶开放复合物（Sasse-Dwight

and Gralla 1989)，后来被用于研究 Pol II 复合物（Wang et al. 1992）。

高锰酸钾修饰单链 DNA 中的胸腺嘧啶残基，但不修饰双链 DNA 中的胸腺嘧啶残基。高锰酸盐已用于鉴定核提取物中启动子上的 Pol II 开放复合物（Wang et al. 1992），或在几种实验生物体中体内启动子上的 Pol II 开放复合物（Giardina and Lis 1993）。在体外，Pol II 对 DNA 的解链是能量依赖性的，因此反应混合物中需加入 ATP。通常使用 dATP，以避免 ATP 偶然掺入转录本的可能性。α-鹅膏蕈碱阻止由核提取物中残存的 NTP 刺激的延伸。通过线性 PCR 引物延伸检测被修饰的胸腺嘧啶。在 PCR 中，*Taq* 酶停留在模板上被修饰的胸腺嘧啶处。产物在聚丙烯酰胺/尿素测序胶上分级分离为一系列带型。高锰酸盐会在双链 DNA 中产生对胸腺嘧啶的低背景修饰。如果开放复合物仅在一小部分 DNA 上形成，就难以检测到背景映衬的信号。因此，利用较宽范围的模板浓度进行初次实验。首选在起始位点附近 T 残基丰富的启动子，因为它们能增强潜在的信号。

关键的是通过引物延伸或一些其他技术定位该起始位点，以确保起始位点区域已知，这种要求在哺乳动物细胞提取物中很重要。但是，酵母中开放复合物和起始位点在定位中有时表现出差异。该反应很像体外转录实验，反应在 30℃进行，含有 DNA 模板、核提取物和激活因子。通常，该反应含有存在于 4～5 kb 质粒上的 10～50 ng 的模板，还含有非特异性载体 DNA。经过一个简单的 30 min 孵育使起始前复合物组装后，加入高锰酸盐，对 DNA 进行修饰（1 min 或更长时间）。然后纯化 DNA，对其进行引物延伸分析，产物在聚丙烯酰胺/尿素测序胶上分级分离。将有无激活因子的反应及有无 ATP 的反应进行比较。建议对 DNA 模板、激活因子和时间进程进行滴定评估。通常，开放复合物的量和转录有很好的相关性。对于带有不同个数位点的模板上的 GAL4-VP16 来说，开放复合物的量和转录量平行（Wang et al. 1992）。另外，通过将 VP16 突变体的转录潜力和形成开放复合物的能力相比较，也观察到了相似的现象（Jiang et al. 1994）。

时线与安排

该方法操作简单，除需要 ^{32}P-末端标记的 DNA 引物外，还需要用于体外转录反应的标准成分。其中，^{32}P-末端标记 DNA 引物定位于转录起始位点下游 60～90 核苷酸处，且与上链互补。探测反应本身和引物延伸可以在一天内完成，第二天对结果进行检测。但需要优化反应条件，这一过程会花费一周或更长的时间。转录水平弱的提取物和启动子会使信号更难检测。高锰酸盐是强氧化剂，孵育时间太长会破坏蛋白质-DNA 相互作用。通常，要对含有或缺失激活因子结合位点的模板上的开放复合物，或添加或不添加激活因子的反应，进行比较。

第 1 天：高锰酸钾探测

第 2 天：数据分析

提纲

高锰酸钾探测（1 天）。

步骤 1：准备缓冲液和试剂（20 min）。
步骤 2：准备反应和进行足迹分析（1 h，30 min）。
步骤 3：进行引物延伸反应（2～3 h）。
步骤 4：在凝胶上分级分离引物延伸产物（3～4 h）。
数据分析

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

α-鹅膏蕈碱（80 μg/μl）〈!〉
乙酸铵（5 mol/L）
β-巯基乙醇〈!〉
缓冲液 D
　　0.2 mmol/L EDTA
　　20% 甘油（V/V）
　　20 mmol/L HEPES-KOH（pH 7.9）〈!〉
　　0.1 mol/L KCl〈!〉
　　临用前加入：
　　　1 mmol/L DTT〈!〉
　　　1 mmol/L PMSF〈!〉
载体 DNA（200 ng/μl）
氯仿〈!〉
dATP（20 mmol/L）
DNA 模板（10 ng/μl）
dNTP（25 mmol/L）
干冰〈!〉
乙醇（80%和 100%）〈!〉
HeLa 细胞核提取物（见方案 12.1）
　　　临用前以缓冲液 D 稀释到 10 mg/ml。
$KMnO_4$（0.37 mol/L）〈!〉
$MgCl_2$（100 mmol/L）〈!〉
矿物油（可选）
引物（^{2}P-标记；0.3 pmol/μl）〈!〉
苯酚〈!〉
蛋白酶 K（10 mg/ml）〈!〉
终止缓冲液
　　10 mmol/L EDTA

0.1% SDS 〈!〉
0.4 mol/L 乙酸钠
50 μg/ml 酵母 tRNA

Taq 酶（4～20 U/μl）
10×*Taq* 酶缓冲液
TE 缓冲液
聚丙烯酰胺/尿素凝胶测序所需其他试剂

设备

离心机
PCR 扩增热循环仪
微量离心管
聚丙烯酰胺/尿素凝胶测序所需其他设备

程序

步骤 1：准备缓冲液和试剂

1. 准备高锰酸钾工作液：立即使用前用 32 μl 水稀释 5 μl 原液，存放在避光瓶中。
2. 准备储存缓冲液。

步骤 2：建立反应及进行足迹分析

1. 准备反应混合液：

核提取物	25 μl
DNA 模板	1.0 μl
载体 DNA	1.0 μl
100 mmol/L $MgCl_2$	3.3 μl
20 mmol/L dATP	1.0 μl
α-鹅膏蕈碱	1.0 μl
H_2O	补至 40 μl

DNA 模板的浓度可以有所变动，但 DNA 总量应保持在 200 ng。

2. 30℃孵育 30 min。
3. 加 5 μl $KMnO_4$ 工作液，30℃孵育 4 min。
4. 加 3 μl β-巯基乙醇淬灭 $KMnO_4$ 反应。
5. 加 100 μl 终止缓冲液和 1 μl 蛋白酶 K，37℃孵育 15 min。
6. 酚抽提至少一次。
7. 酚：氯仿（1：1）抽提，乙醇沉淀水相。14 000 r/min 离心 10 min。
8. 用 80%乙醇洗涤沉淀，空气干燥。
9. 将沉淀重悬于 60 μl TE 中。

步骤 3：进行引物延伸反应

当使用引物延伸检测产物时，重要的是要有一个模拟处理的 DNA 或单独在提取物中孵育的 DNA，并将其作为对照。引物延伸有时对未经处理 DNA 中的自然暂停序列敏感。

1. 准备引物延伸反应：

DNA 产物（来自步骤 2：9）	20 μl
10×*Taq* 酶缓冲液	5.0 μl
25 mmol/L dNTP	0.5 μl
^{32}P-标记的引物	1.0 μl
Taq 酶（1～5 U）	0.25 μl
100 mmol/L $MgCl_2$	1.5 μl
H_2O	补至 50 μl

正如在许多 *Taq* 酶依赖性引物延伸反应中那样，镁离子浓度必须依经验而定。

2. 矿物油覆盖反应（否则使用具有绝热室的热循环仪）。

3. 热循环仪设置如下：90℃，1 min；55℃，2 min；72℃，2 min。重复 30 个循环。

4. 弃上清，加入 33 μl 5 mol/L 乙酸铵。再加入 250 μl 乙醇，充分混匀，冰上孵育 10 min。

5. 在微量离心机中 14 000r/min 离心 10 min，沉淀 DNA。用 80%乙醇洗涤沉淀，空气干燥。

6. 将沉淀重悬于 4 μl 甲酰胺测序染料中，然后与以引物延伸反应所使用的相同引物产生的双脱氧测序梯带并排在 8%测序胶上分级分离。

步骤 4：数据分析

利用分子质量标准物，鉴定开放复合物的假定位置。开放复合物应在起始位点处形成。凝胶上活化泳道和非活化泳道之间应显示出明显差异。有时，由于哺乳动物核提取物的内源性核苷酸，在 ATP 不存在的条件下也能观察到开放复合物。

附加说明

1. 该方法可以用末端标记的 DNA 取代引物延伸来检测被修饰的胸腺嘧啶（T）。在这种情况下，和 Maxam-Gilbert 测序一样，首先用哌啶处理反应产物，或不经处理也经常发生链剪切，产物可以直接上样到凝胶上。

2. 当利用引物延伸法检测产物时，重要的是用模拟处理的 DNA 或单独在提取物中孵育的 DNA 作为对照。引物延伸有时对未经处理的 DNA 中的天然暂停序列敏感。

3. 该方法也可用于研究延伸和启动子逃脱中的早期事件。

疑难解答

无特异信号

观察到了高锰酸盐依赖性带型，但在激活因子和 ATP 存在与否的情况下，看不到

开放复合物信号。

可能原因：核提取物无活性，或模板不含位于起始位点附近的可及性胸腺嘧啶。

解决办法：在提取物中分析转录激活作用，以确保活性。降低模板浓度通常能改善信噪比。通过诱变往模板中加入更多的胸腺嘧啶，以提高灵敏度。

无修饰信号

可能原因：提取物中过多的核苷酸会导致延伸并不能观察到信号。

解决办法：确保加入了 α-鹅膏蕈碱，或加入三磷酸腺苷双磷酸酶，以消除 ATP。

观察不到高锰酸盐依赖性修饰

可能原因：高锰酸盐可能已经变质。

解决办法：使用较高浓度的新鲜高锰酸盐，或稍微延长孵育时间。

方案 12.8 DNA 结合 TFIID 的镁-琼脂糖 EMSA

因为通用转录因子 TFIID 与核心启动子的结合是转录复合物组装中的成核步骤，所以 TFIID 是调控的关键靶标。当在体外将启动子上的 TFIID 调到极限浓度时，许多真核激活因子都能刺激 TFIID 的募集。传统上，通过 DNase I 足迹法研究激活因子刺激的 TFIID 募集。但是研究已经表明，镁-琼脂糖凝胶能够分离包含 TFIID、TFIIA（DA 复合物）和 TFIIB（DAB 复合物）的大复合物，并能够定量地测定激活因子如何刺激这些复合物的组装。EMSA 的优点是反应能够在 TFIID 足迹可能不能观察到的亚饱和条件下进行（Lieberman and Berk 1994；Zerby and Lieberman 1997）。通常将激活因子与 ^{32}P 标记的 DNA 模板、从 *E. coli* 中纯化的重组 TFIIA 及免疫纯化的 TFIID 一起孵育。TFIID 是来自 Berk 及其同事构建的一个特殊株系的 HA 表位标记型蛋白（Zhou et al. 1992），但 Roeder 及其同事（Chiang et al. 1993）获得的 FLAG 标记 TFIID 也有效。

包含 TFIID 的 DNA-蛋白质复合物在镁-琼脂糖凝胶上的迁移率下降有两个原因，包括：①TBP 亚基诱导导致迁移率改变的 DNA 尖锐弯曲；②TBP-TAF_{II}复合物的块头大。通过比较激活因子存在与否时 TFIID 在大范围浓度下的结合，就可以评价激活因子是否与 TBP 或一种 TAF_{II}发生相互作用。也可以相继加入其他因子（如 TFIIA 和 TFIIB），以量化它们对转录复合物组装的贡献。

该方法灵敏且信息量大，但是与多数生化分析一样，其结果应和体外转录研究（即动力学分析）相结合，以评价转录过程中 DA 复合物的生理相关性。我们发现，DA 复合物的组装和基因活化水平、激活因子刺激的开放复合物的形成密切相关（Chi and carey 1993；Chi et al. 1995）。

时线与安排

该方案相对简单，但需要不太容易得到的特殊试剂，包括通用转录因子 TFIIA、

TFIIB 和 TFIID。在成功地使用该项技术之前，要花费大量时间纯化并表征这些试剂。TFIIA 和 TFIIB 能够以重组蛋白的形式得到。但是 TFIID 难以获得，必须具备大规模培养设备，以培养含有表位标记 TFIID 的菌株。一个日益普遍使用的方法是，利用来自 *E. coli* 的 TBP 和在杆状病毒系统中过量表达的 TAF_{II} 构建重组 TBP-TAF_{II} 亚复合物（Chen and Tjian 1996）。

理想情况下，结合反应中除含有各种蛋白质外，还含有 10 ng 约 200 bp 的 ^{32}P 标记 DNA 模板。有时，加入 0.1 % Tween-20 以防止聚集。通常，30℃孵育 1 h 后，将混合物加到 10 cm 长，用 45 mmol/L Tris 碱、45 mmol/L 硼酸和 5 mmol/L 乙酸镁配制的 1.4%水平琼脂糖凝胶（约 100 ml）上；在同样的缓冲液中以 50 V 电泳 3 h，使凝胶在 Whatman DE81 滤纸上干燥，通过 XAR-5 光片曝光进行放射自显影或磷屏成像分析（Kodak）。为了得到一致性的结果，琼脂糖应为低 EEO 种类。

第 1 天：^{32}P 标记启动子片段（见方案 12.6）。

第 2 天：镁-琼脂糖凝胶电泳。

第 3 天：数据分析。

提纲

以 [γ-^{32}P] ATP 标记的启动子片段（见方案 12.6）。

镁-琼脂糖凝胶电泳（1 天）。

步骤 1：配制缓冲液（30 min）。

步骤 2：灌制镁-琼脂糖凝胶（30 min）。

步骤 3：建立结合反应（1 h）。

步骤 4：电泳（3 h）。

步骤 5：放射自显影（1 h）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

琼脂糖 [低电渗（EEO）；Fisher，BP160-500]

为了结果更加一致，应采用低 EEO 琼脂糖。

10×凝胶缓冲液

0.45 mol/L 硼酸〈!〉

50 mmol/L 乙酸镁

0.45 mol/L Tris 碱〈!〉

临用前稀释至 1×。

β-巯基乙醇（14 mol/L）〈!〉

牛血清白蛋白（BSA，50 mg/ml）

缓冲液 D

0.2 mmol/L EDTA
20%甘油（V/V）
20 mmol/L HEPES-KOH（pH7.9）〈!〉
0.1 mol/L KCl〈!〉
dG∶dC（1.0 mg/ml）
DNA模板（100 ng/μl），γ-^{32}P末端标记（10 000 cpm/μl）〈!〉
启动子片段应约200 bp。
$MgCl_2$（0.1mol/L）〈!〉
PMSF（100 mmol/L）〈!〉
重组TFIIA（20 ng/μl）
重组TFIIB（160 ng/μl）
表位标记TFIID（经免疫纯化，200 ng/μl）
所采用的所有因子都应在使用前用缓冲液D透析。

设备

DEAE纤维素纸（如DE81；Whatman，3658-917）
电泳单元
滤纸（Whatman）
磷屏成像仪暗盒（可选）
保鲜膜
硅化微量离心管（0.5 ml）
真空干燥机
X光片（如XAR-5，Kodak）（可选）

程序

步骤1：配制缓冲液

配制10×凝胶缓冲液和缓冲液D。

步骤2：灌制镁-琼脂糖凝胶

1. 选用低EEO的琼脂糖，用1×凝胶缓冲液配制1.4%镁-琼脂糖凝胶，使混合液稍冷却。
2. 倒制小胶（即10 cm长、8 cm宽、0.5 cm厚），避免凝胶结块。
凝胶一般尺寸很小，应该短时间电泳。为减少解离，有时在冷室中电泳。

步骤3：建立结合反应

典型的反应将包含激活因子和TFIID，并补加TFIIA或TFIIB。下述反应条件包括全部三种通用转录因子，但理想的用量应该在广范围上进行滴定。

1. 在离心管中于于冰上准备反应混合物，组分如下：

组分	体积
TFIID	2.0 μl
TFIIA	1.0 μl
TFIIB	1.0 μl
激活因子	1.0 μl
^{32}P-标记探针	0.1 μl
dG∶dC	0.2 μl
BSA	0.1 μl
β-巯基乙醇	0.05 μl
0.1 mol/L $MgCl_2$	0.75 μl
PMSF	0.05 μl
H_2O	3.35 μl
缓冲液 D	3.0 μl
总体积	12.6 μl

适当乘以每种试剂用量，准备一批混合液以适应多个反应。请注意，可单独用 TFIID 分析结合，或 TFIID 与 TFIIA 和 TFIIB 一起分析结合。因此，如果加入任何转录因子，反应体积将发生改变，通过加入带有缓冲液 D 的稀释剂调整反应，保持氯化钾的终浓度为 60 mmol/L，缓冲液 D 的终浓度为 0.6×。

2. 将混合液于 30℃孵育 15～60 min。
   如果空间和试剂允许，建议至少取两个时间点。

步骤 4：电泳

1. 将反应混合物轻轻加入凝胶（来自步骤 2）加样孔。注意让混合物在加样孔内均匀沉积。缓冲液 D 中的甘油应足以使样品在加样孔内均匀沉积。
2. 在旁边的泳道上上样一个标准的凝胶染料混合物。
3. 在 1×凝胶缓冲液中 50 V 室温电泳凝胶 3 h。
   溴酚蓝应保留在凝胶上。对于更短的凝胶，使用更短的电泳时间。凝胶也可在冷室内电泳，以减少解离。
4. 剪一片凝胶板大小的 DEAE 纤维素纸。
5. 将胶放置在纸上，用保鲜膜覆盖凝胶，下面垫两片滤纸。
6. 65℃真空干燥凝胶 1 h（或直至完全干燥）。

步骤 5：放射自显影或磷屏成像分析

1. 将干凝胶转移到 X 光片上或磷屏成像盒中放射自显影。
2. 分析放射自显影确定电泳迁移率。

- 理想的情况下，游离探针恰好迁移至溴酚蓝染料前沿。
- 单独激活因子的加入一般仅引起迁移率的微小变化（除非激活因子改变了 DNA 的形状）。

- 单独的 DA 复合物比单独的激活因子应产生更大的泳动迁移率。
- 激活因子-DA 复合物的迁移可能不会明显地高于单独的 DA 复合物。
- DA-激活因子复合物与单独的激活因子间的相对比例应大于单独的 DA 复合物与游离探针的比例。
- 使用这些复合物的比例（经磷屏成像仪确定）仔细定量激活因子的效应。

附加说明

1. 为了观察到激活因子的效应，必须在较宽的范围内对 TFIID 进行滴定。TFIID 在低浓度时，几乎看不到对 ZEBRA 引起的募集有何影响（Chi et al. 1995），而 TFIID 在高浓度时，没有加入激活因子 DNA 就被饱和。关键是要进行剂量应答滴定，以鉴定相关的浓度窗口。

2. TFIIA、TFIIB 和几乎可以肯定的其他因子都会影响激活因子募集复合物的能力。因为所有这些转录复合物的组装反应都有一定程度的协同性，一个与 TFIIB 发生强相互作用的激活因子会通过与 TFIIB 的结合、TFIIB 与 TFIID 的结合而间接募集 TFIID。同样，重要的是要通过滴定定量这些额外的贡献。

3. 由于一些未知原因，含有 TBP 的复合物在这些凝胶中的迁移率有时和 TFIID 的迁移率非常相似。虽然 TBP 复合物通常不受激活因子影响，但它可用于研究缺乏 TAF_{II}时激活因子-TFIIA 和激活因子-TFIIB 的相互作用对 TBP 募集的影响。

4. 凝胶系统要经得起许多不同条件的检验。有时难于将一个复合物和另一个复合物（即 DA 和 DAB 复合物）分开。滴定乙酸镁的浓度有时有助于辨别两种复合物。另外，凝胶可以在多种温度条件下电泳，温度能改变各复合物的相对迁移率。结合反应自身也能忍受盐浓度的变化。

疑难解答

无 DA 复合物或激活因子结合

可能原因：有理由认为在有些分析条件下某些激活因子可能不与启动子 DNA 稳定地结合，但 DA 不结合是不可能的。

解决办法：如果未观察到 TFIID 结合，就应加入更多的 TFIID 组分。前面已经提到，TFIID 对不同核心启动子的亲和力可能有很大的变动，该方法倾向于在窄的 TFIID 浓度窗口内最为有效。因此，未观察到结合时，简单提高 TFIID 浓度是可取的。

拖尾凝胶

可能原因：有时，亲和纯化的 TFIID 制剂中仍含有一些蛋白质 A-琼脂糖凝胶微珠，它们会将复合物滞留在上样孔内。

解决办法：为了避免这种情况，使用前将制备的 TFIID 在离心机上离心，并小心地将上清转移至一新的硅化管中。低纯度的 TFIID 也会导致复合物滞留在上样孔内。TFIID 免疫纯化程序中多漂洗几次能够去除杂质，以缓解这一问题。另外，要注意确

保使用低 EEO 琼脂糖。检查蛋白质的纯度，加入 0.05 % NP-40。

激活因子对 DA 无刺激

如果在一系列条件下观察到 DA 复合物，并且 DA 复合物以剂量应答方式积累，但激活因子在此期间没有刺激其形成，就可以推断激活因子未与 TFIID 直接相互作用。比如，含有完整 VP16 激活结构域的 GAL4 衍生物刺激了 DA 复合物的形成，而仅含有 VP16 激活结构域氨基端一半的相似衍生物不能刺激 DA 复合物的形成，但是二者在 HeLa 体外系统中都能很好地激活转录。这类数据说明在激活因子和通用机器之间存在着其他对体外基因活化更重要的相互作用。

参考文献

Akoulitchev S., Chuikov S., and Reinberg D. 2000. TFIIH is negatively regulated by cdk8-containing mediator complexes. *Nature* 407: 102-106.

An W., Kim J., and Roeder R. G. 2004. Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. *Cell* 117: 735-748.

Andrau J. C., van de Pasch L., Lijnzaad P., Bijma T., Koerkamp M. G., van de Peppel J., Werner M., and Hoistege F. C. 2006. Genomewide location of the coactivator mediator: Binding without activation and transient Cdk8 interaction on DNA. *Mol. Cell* 22: 179-192.

Aoyagi N. and Wassarman D. A. 2001. Developmental and transcriptional consequences of mutations in *Drosophila* $TAF_{II}60$. *Mol. Cell. Biol.* 21: 6808-6819.

Apone L. M., Virbasius C. M., Reese J. C., and Green M. R. 1996. Yeast $TAF_{II}90$ is required for cell-cycle progression through G2/M but not for general transcription activation. *Genes Dev.* 10: 2368-2380.

Araujo S. J., Tirode F., Coin F., Pospiech H., Syvaoja J. E., Stucki M., Hubscher U., Egly T. M., and Wood R. D. 2000. Nucleotide excision repair of DNA with recombinant human proteins: Definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes Dev.* 14: 349-359.

Arias J. A. and Dynan W. S. 1989. Promoter-dependent transcription by RNA polymerase II using immobilized enzyme complexes. *J. Biol. Chem.* 264: 3223-3229.

Armache K. J., Kettenberger H., and Cramer P. 2003. Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc. Natl. Acad. Sci.* 100: 6964-6968.

Asada S., Choi Y., Yamada M., Wang S. C., Hung M. C., Qin J., and Uesugi M. 2002. External control of Her2 expression and cancer cell growth by targeting a Ras-linked coactivator. *Proc. Natl. Acad. Sci.* 99: 12747-12752.

Aso T., Vasavada H. A., Kawaguchi T., Germino F. J., Ganguly S., Kitajima S., Weissman S. M., and Yasu kochi Y. 1992. Characterization of cDNA for the large subunit of the transcription initiation factor TFIIF. *Nature* 355: 461-464.

Asturias F. J., Jiang Y. W., Myers L. C., Gustafsson C. M., and Kornberg R. D. 1999. Conserved structures of Mediator and RNA polymerase II holoenzyme. *Science* 283: 985-987.

Baek H. J., Kang Y. K., and Roeder R. G. 2006. Human mediator enhances basal transcription by facilitating recruitment of transcription factor IIβ during preinitiation complex assembly. *J. Biol. Chem.* 281: 15172-15181.

Baek H. J., Malik S., Qin J., and Roeder R. G. 2002. Requirement of TRAP/Mediator for both activator-independent and activator-dependent transcription in conjunction with TFIID-associated $TAF_{II}s$. *Mol. Cell. Biol.* 22: 2842-2852.

Balciunas D., Hallberg M., Bjorklund S., and Ronne H. 2003. Functional interactions within yeast Mediator and

evidence of differential subunit modifications. *J. Biol. Chem.* 278: 383 1-3839.

Banerjee S., Zhang Y., Ali S., Bhuiyan M., Wang Z., Chiao P. J., Philip P. A., Abbruzzese J., and Sarkar F. H. 2005. Molecular evidence for increased antitumor activity of gemcitabine by genistein in vitro and in vivo using an orthotopic model of pancreatic cancer. *Cancer Res.* 65: 9064-9072.

Bartholomew B., Kassavetis G. A., and Geiduschek E. P. 1991. Two components of *Saccharomyces cerevisiae* transcription factor IIIB (TFIIIB) are stereospecifically located upstream of a tRNA gene and interact with the second-largest subunit of TFIIIC. *Mol. Cell. Biol.* 11: 5181-5189.

Bartholomew B., Durkovich D., Kassavetis G. A., and Geiduschek E. P. 1993. Orientation and topography of RNA polymerase III in transcription complexes. *Mol. Cell. Biol.* 13: 942-952.

Barton M. C. and Emerson B. M. 1996. Regulated gene expression in reconstituted chromatin and synthetic nuclei. *Methods Enzymol.* 274: 299-312.

Baumli S., Hoeppner S., and Cramer P. 2005. A conserved Mediator hinge revealed in the structure of the MED7MED21 (Med7Srb7) heterodimer. *J. Biol. Chem.* 280: 18171-18178.

Becker P. B., Tsukiyama T., and Wu C. 1994. Chromatin assembly extracts from *Drosophila* embryos. *Methods Cell Biol.* 44: 207-223.

Belakavadi M. and Fondell J. D. 2006. Role of the Mediator complex in nuclear hormone receptor signaling. *Rev. Physiol. Biochem. Pharniacol.* 156: 23-43.

Bell B., Scheer E., and Tora L. 2001. Identification of $hTAF_{II}80\ \delta$ links apoptotic signaling pathways to transcription factor TFIID function. *Mol. Cell* 8: 591-600.

Bentley D. L. 1995. Regulation of transcriptional elongation by RNA polymerase II. *Curr. Opin. Genet. Dev.* 5: 210-216.

Berk A. J. 2005. Recent lessons in gene expression, cell cycle control, and cell biology from adenovirus. *Oncogene* 24: 7673-7685.

Bernstein R., Dejong J., and Roeder R. G. 1994. Characterization of the highly conserved TFIIA small subunit from *Drosophila melanogaster*. *J. Biol. Chem.* 269: 24361-24366.

Bertolotti A., Lutz Y., Heard D. J., Chambon P., and Tora L. 1996. $hTAF_{II}68$, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *EMBO J.* 15: 5022-5031.

Beve J., Hu G. Z., Myers L. C., Balciunas D., Werngren O., Hultenby K., Wibom R., Ronne H., and Gustafsson C. M. 2005. The structural and functional role of MedS in the yeast Mediator tail module. *J. Biol. Chem.* 280: 41366-41372.

Birkenmeier E. H., Brown D. D., and Jordan E. 1978. A nuclear extract of *Xenopus laevis* oocytes that accurately transcribes 5S RNA genes. *Cell* 15: 1077-1086.

BjBrklund S. and Gustafsson C. M. 2005. The yeast Mediator complex and its regulation. *Trends Biochem. Sci.* 30: 240-244.

Black J. C., Choi J. E., Lombardo S. R., and Carey M. 2006. A mechanism for coordinating chromatin modification and preinitiation complex assembly. *Mol. Cell* 23: 809-818.

Blazek E., Mittler G., and Meisterernst M. 2005. The mediator of RNA polymerase II. *Chromosoma* 113: 399-408.

Boeger H., Bushnell D. A., Davis R., Griesenbeck J., Lorch Y., Strattan J. S., Westover K. D., and Kornberg R. D. 2005. Structural basis of eukaryotic gene transcription. *FEBS Lett.* 579: 899-903.

Boube M., Joulia L., Cribbs D. L., and Bourbon H. M. 2002. Evidence for a mediator of RNA polymerase II transcriptional regulation conserved from yeast to man. *Cell* 110: 143-151.

Boube M., Faucher C., Joulia L., Cribbs D. L., and Bourbon H. M. 2000. *Drosophila* homologs of transcriptional mediator complex subunits are required for adult cell and segment identity specification. *Genes Dev.* 14:

2906-29 17.

Bour G., Gaillard E., Bruck N., Lalevee S., Plassat J. L., Busso D., Samama J. P., and Rochette-Egly C. 2005. Cyclin H binding to the RARce activation function (AF) -2 domain directs phosphorylation of the AF-1 domain by cyclin-dependent kinase 7. *Proc. Natl. Acad. Sci.* 102: 16608-16613.

Bradsher J., Coin F., and Egly J. M. 2000. Distinct roles for the helicases of TFIIH in transcript initiation and promoter escape. *J. Biol. Chem.* 275: 2532-2538.

Brower C. S., Sato S., Tomomori-Sato C., Kamura T., Pause A., Stearman R., Klausner R. D., Malik S., Lane W. S., Sorokina I., et al. 2002. Mammalian mediator subunit mMED8 is an Elongin BC-interacting protein that can assemble with Cul2 and Rbxl to reconstitute a ubiquitin ligase. *Proc. Natl. Acad. Sci.* 99: 10353-10358.

Bulger M., Ito I., Kamakaka R. T., and Kadonaga J. T. 1995. Assembly of regularly spaced nucleosome arrays by *Drosophila* chromatin assembly factor 1 and a 56-kDa histone-binding protein. *Proc. Natl. Acad. Sci.* 92: 11726-11730.

Bunick D., Zandomeni R., Ackerman S., and Weinmann R. 1982. Mechanism of RNA polymerase II-Specific initiation of transcription in vitro: ATP requirement and uncapped runoff transcripts. *Cell* 29: 877-886.

Buratowski S., Hahn S., Guarente L., and Sharp P. A. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* 56: 549-561.

Burke T. W. and Kadonaga J. T. 1997. The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by$TAF_{II}60$ of *Drosophila*. *Genes Dev.* 11: 3020-3031.

Bushnell D. A. and Kornberg R. D. 2003. Complete, 12-subunit RNA polymerase II at 4. 1Å resolution: Implications for the initiation of transcription. *Proc. Natl. Acad. Sci.* 100: 6969-6973.

Bushnell D. A., Westover K. D., Davis RE., and Kornberg R. D. 2004. Structural basis of transcription: An RNA polymerase II-TFIIB cocrystal at 4. 5 angstroms. *Science* 303: 983-988.

Cai H. and Lose D. S. 1987. Transcription initiation by RNA polymerase II in vitro. Properties of preinitiation, initiation, and elongation complexes. *J. Biol. Chem.* 262: 298-304.

Cal K., Itoh Y., and Khorana H. G. 2001. Mapping of contact sites in complex formation between transducin and light-activated rhodopsin by covalent crosslinking: Use of a photoactivatable reagent. *Proc. Natl. Acad. Sci.* 98: 4877-4882.

Carey M. 1998. The enhanceosome and transcriptional synergy. *Cell* 92: 5-8.

Carey M., Leatherwood J., and Ptashne M. 1990a. A potent GAL4 derivative activates transcription at a distance in vitro. *Science* 247: 710-712.

Carey M., Lin Y. S., Green M. R., and Ptashne M. 1990b. A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. *Nature* 345: 361-364.

Carey M. F., Gerrard S. P., and Cozzarelli N. R. 1986. Analysis of RNA polymerase III transcription. complexes by gel filtration. *J. Biol. Chem.* 261: 4309-4317.

Caron C., Mengus G., Dubrowskaya V., Roisin A., Davidson I., and Jalinot p. 1997. Human $TAF_{II}28$ interacts with the human T cell leukemia virus type I Tax transactivator and promotes its transcriptional activity. *Proc. Natl. Acad. Sci.* 94: 3662-3667.

Carrozza M. J., Hassan A. H., and Workman J. L. 2003. Assay of activator recruitment of chromatin-modifying complexes. *Methods Enzymol.* 371: 536-544.

Chadick J. Z. and Asturias F. J. 2005. Structure of eukaryotic Mediator complexes. *Trends Biochem. Sci.* 30: 264-271.

Chalkley G. E. and Verrijzer C. P. 1999. DNA binding site selection by RNA polymerase II TAFs: A $TAF_{II}250$-$TAF_{II}150$ complex recognizes the initiator. *EMBO J.* 18: 4835-4845.

Chalut C., Lang C., and Egly J. M. 1994. Expression in *Escherichia coli*: Production and purification of both sub-

units of the human general transcription factor TFIIE. *Protein Expr. Purif.* 5: 458-467.

Chen B. S. and Hampsey M. 2002. Transcription activation: Unveiling the essential nature of TFIID. *Curr. Biol.* 12: R620-622.

Chen H. T. and Hahn S. 2004. Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: A model for the structure of the PIC. *Cell* 119: 169-180.

Chen J., Larochelle S., Li X., and Suter B. 2003. Xpd/Ercc2 regulates CAK activity and mitotic progression. *Nature* 424: 228-232.

Chen J. L. and Tjian R. 1996. Reconstitution of TATA-binding protein-associated factor/TATA-binding protein complexes for in vitro transcription. *Methods Enzymol.* 273: 208-217.

Chen J. L., Attardi L. D., Verrijzer C. P., Yokomori K., and Tjian R. 1994. Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. *Cell* 79: 93-105.

Chen W., Rogatsky I., and Garabedian M. J. 2006. MED14 and MED1 differentially regulate target-specific gene activation by the glucocorticoid receptor. *Mol. Endocrinol.* 20: 560-572.

Chen Y., Ebright Y. W., and Ebright R. H. 1994. Identification of the target of a transcription activator protein by protein-protein photocrosslinking. *Science* 265: 90-92.

Chen Z. and Manley J. L. 2003. In vivo functional analysis of the histone 3-like TAF9 and a TAF9-related factor, TAF9L. *J. Biol. Chem.* 278: 35172-35183.

Chi T. and Carey M. 1993. The ZEBRA activation domain: Modular organization and mechanism of action. *Mol. Cell. Biol.* 13: 7045-7055.

Chi T. and Carey M. 1996. Assembly of the isomerized TFIIA-TFIID-TATA ternary complex is necessary and sufficient for gene activation. *Genes Dev.* 10: 2540-2550.

Chi T., Lieberman P., Ellwood K., and Carey M. 1995. A general mechanism for transcriptional synergy by eukaryotic activators. *Nature* 377: 254-257.

Chiang CM. and Roeder R. G. 1995. Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. *Science* 267: 531-536.

Chiang C. M., Ge H., Wang Z., Hoffmann A., and Roeder R. G. 1993. Unique TATA-binding protein-containing complexes and cofactors involved in transcription by RNA polymerases II and III. *EMBO J.* 12: 2749-2762.

Choy B. and Green M. R. 1993. Eukaryotic activators function during multiple steps of preinitiation complex assembly. *Nature* 366: 531-536.

Chung W. H., Craighead IL., Chang W. H., Ezeokonkwo C., BareketSamish A., Kornberg R. D., and Asturias F. J. 2003. RNA polymerase II/TFIIF structure and conserved organization of the initiation complex. *Mol. Cell* 12: 1003-1013.

Coin F., Auriol J., Tapias A., Clivio P., Vermeulen W., and Egly J. M. 2004. Phosphorylation of XPB helicase regulates TFIIH nucleotide excision repair activity. *EMBO J.* 23: 4835-4846.

Conaway J. W. and Conaway R. C. 1989. A multisubunit transcription factor essential for accurate initiation by RNA polymerase II. *J. Biol. Chem.* 264: 2357-2362.

Conaway J. W., Bond M. W., and Conaway R. C. 1987. An RNA polymerase II transcription system from rat liver. Purification of an essential component. *J. Biol. Chem.* 262: 8293-8297.

Conaway J. W., Reines D., and Conaway R. C. 1990. Transcription initiated by RNA polymerase II and purified transcription factors from liver. Cooperative action of transcription factors t and e in initial complex formation. *J. Biol. Chem.* 265: 7552-7558.

Conaway J. W., Hanley J. P., Garrett K. P., and Conaway R. C. 1991. Transcription initiated by RNA polymerase II and transcription factors from liver. Structure and action of transcription factors ε and τ. *J. Biol. Chem.* 266: 7804-7811.

Conaway R. C. and Conaway J. W. 1989. An RNA polymerase II transcription factor has an associated DNA-depend-

ent ATPase (dATPase) activity strongly stimulated by the TATA region of promoters. *Proc. Natl. Acad. Sci.* 86: 7356-7360.

Conaway R. C., Reines D., Garrett K. P., Powell W., and Conaway J. W. 1996. Purification of RNA polymerase II general transcription factors from rat liver. *Methods Enzymol*. 273: 194-207.

Conaway R. C., Sato S., Tomomori-Sato C., Yao T., and Conaway J. W. 2005. The mammalian Mediator complex and its role in transcriptional regulation. *Trends Biochem. Sci.* 30: 250-255.

Cramer P., Bushnell D. A., and Kornberg R. D. 2001. Structural basis of transcription: RNA polymerase II at 2. 8 angstrom resolution. *Science* 292: 1863-1876.

Cramer P., Bushnell D. A., Fu J., Gnatt A. L, Maier-Davis B., Thompson N. E., Burgess R. R., Edwards A. M., David P. R., and Kornberg R. D. 2000. Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 288: 640-649.

Dahan O. and Kupiec M. 2004. The *Saccharomyces cerevisiae* gene *CDC40/PRP1* 7 controls cell cycle progression through splicing of the *ANCI* gene. *Nucleic Acids Res*. 32: 2529-2540.

DeJong J. and Roeder R. G. 1993. A single cDNA, hTFIIAJz, encodes both the p35 and p19 subunits of human TFIIA. *Genes Dev*. 7: 2220-2234.

Dejong J., Bernstein R., and Roeder R. G. 1995. Human general transcription factor TFIIA: Characterization of a cDNA encoding the small subunit and requirement for basal and activated transcription. *Proc. Natl. Acad. Sci.* 92: 3313-3317.

Di Lello P., Jenkins L. M., Jones T. N., Nguyen B. D., Hara T., Yamaguchi H., Dikeakos J. D., Appella E., Legault P., and Omichinski J. G. 2006. Structure of the Tfbl/p53 complex: Insights into the interaction between the p62/Tfbl subunit of TFIIH and the activation domain of p53. *Mol. Cell* 22: 731-740.

Dignam J. D., Lebovitz R. M., and Roeder R. G. 1983a. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res*. 11: 1475-1489.

Dignam J. D., Martin P. L., Shastry B. S., and Roeder R. G. 1983b. Eukaryotic gene transcription with purified components. *Methods Enzymol*. 101: 582-598.

Douziech M., Coin F., Chipoulet J. M., Arai Y., Ohkuma Y., Egly J. M., and Coulombe B. 2000. Mechanism of promoter melting by the xeroderma pigmentosum complementation group B helicase of transcription factor IIH revealed by protein-DNA photocrosslinking. *Mol. Cell. Biol.* 20: 8168-8177.

Dubaele S., Proietti De Santis L., Bienstock R. J., Keriel A., Stefanjni M., Van Houten B., and Egly J. M. 2003. Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients. *Mol. Cell* 11: 1635-1646.

Dubrovskaya V., Lavigne A. C., Davidson I., Acker J., Staub A., and Tora L. 1996. Distinct domains of $hTAF_{II}100$ are required for functional interaction with transcription factor TFIIβ (RAP30) and incorporation into the TFIID complex. *EMBOJ*. 15: 3702-3712.

Dürr K., Holzschuh J., Filippi A., EttI AK., Ryu S., Shepherd IT., and Driever W. 2006. Differential roles of transcriptional Mediator complex subunits Crsp34/Med27, Crsp150/Med14, and Trap 100/Med24 during zebrafish retinal development. *Genetics* 174: 693-705.

Dvir A., Conaway J. W., and Conaway R. C. 2001. Mechanism of transcription initiation and promoter escape by RNA polymerase II. *Curr. Opin. Genet. Dev.* 11: 209-214.

Dvir A., Conaway R. C., and Conaway J. W. 1996. Promoter escape by RNA polymerase II. A role for an ATP cofactor in suppression of arrest by polymerase at promoter-proximal Sites. *J. Biol. Chem.* 271: 23352-23356.

Dvir A., Conaway R. C., and Conaway J. W. 1997a. A role for TFIIH in controlling the activity of early RNA polymerase II elongation complexes. *Proc. Natl. Acad. Sci.* 94: 9006-9010.

Dvir A., Tan S., Conaway J. W., and Conaway R. C. 1997b. Promoter escape by RNA polymerase II. Formation of an escape-competent transcriptional intermediate isa prerequisite for exit of polymerase from the promoter. *J.*

Biol. Chem. 272：28175-28178.

Dynan W. S. and Tjian R. 1983. Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. *Cell* 32：669-680.

Egly J. M. 2001. The 14th Datta Lecture. TFIIH：From transcription to clinic. *FEBS Lett.* 498：124-128.

Ellwood K.，Chi T.，Huang W.，Mitsouras K.，and Carey M. 1998. Cooperative assembly of RNA polymerase II transcription complexes. *Cold Spring Harbor Syrup. Quant. Biol.* 63：253-262.

Evangelista C. C.，Jr.，Rodriguez Torres A. M.，Limbach M. P.，and Zitomer R. S. 1996. Rox3 and Rtsl function in the global stress response pathway in baker' s yeast. *Genetics* 142：1083-1093.

Falender A. E.，Freiman R. N.，Geles K. G.，Lo K. C.，Hwang K.，Lamb D. J.，Morris P. L.，Tjian R.，and Richards J. S. 2005. Maintenance of spermatogenesis requires TAF4b，a gonad-specific subunit of TFIID. *Genes Dev.* 19：794-803.

Fan L.，Arvai A. S.，Cooper P. K.，Iwai S.，Hanaoka F.，and Tamer J. A. 2006. Conserved XPB core structure and motifs for DNA unwinding：Implications for pathway selection of transcription or excision repair. *Mol. Cell* 22：27-37.

Feaver W. J.，Huang W.，and Friedberg E. C. 1999. The TFB4 subunit of yeast TFIIH is required for both nucleotide excision repair and RNA polymerase II transcription. *J. Biol. Chem.* 274：29564-29567.

Feaver W. J.，Henry N. L.，Bushnell D. A.，Sayre M. H.，Brickner J. H.，Gileadi O.，and Kornberg R. D. 1994. Yeast TFIIE. Cloning，expression，and homology to vertebrate proteins. *J. Biol. Chem.* 269：27549-27553.

Feaver W. J.，Henry N. L.，Wang Z.，Wu X.，Svejstrup J. Q.，Bushnell D. A.，Friedberg E. C.，and Kornberg RD. 1997. Genes for Tfb2，Tfb3，and Tfb4 subunits of yeast transcription/repair factor IIH. Homology to human cyclin-dependent kinase activating kinase and IIH subunits. *J. Biol. Chem.* 272：19319-19327.

Finkelstein A.，Kostrub C. F.，Li J.，Chavez D. P.，Wang B. Q.，Fang S. M.，Greenblatt J.，and Burton Z. F. 1992. A cDNA encoding RAP74，a general initiation factor for transcription by RNA polymerase II. *Nature* 355：464-467.

Fischer L.，Gerard M.，Chalut C.，Lutz Y.，Humbert S.，Kanno M.，Chambon P.，and Egly J. M. 1992. Cloning of the 62-kilodalton component of basic transcription factor BTF2. *Science* 257：1392-1395.

Fishburn J.，Mohibullah N.，and Hahn S. 2005. Function of a eukaryotic transcription activator during the transcription cycle. *Mol. Cell* 18：369-378.

Fisher R. P. 2005. Secrets of a double agent：CDK7 in cell-cycle control and transcription. *J. Cell Sci.* 118：5171-5180.

Fisher R. P. and Morgan DO. 1994. A novel cyclin associates with MO15/CDK7 to form the CDK-activating kinase. *Cell* 78：713-724.

Flanagan P. M.，Kelleher R. J.，3rd，Sayre M. H.，Tschochner H.，and Kornberg R. D. 1991. A mediator required for activation of RNA polymerase II transcription in vitro. *Nature* 350：436-438.

Forget D.，Langelier M. E，Therien C.，Trinh V.，and Coulombe B. 2004. Photo-crosslinking of a purified preinitiation complex reveals central roles for the RNA polymerase II mobile clamp and TFIIE in initiation mechanisms. *Mol. Cell. Biol.* 24：1122-1131.

Foulds C. E. and Hawley D. K. 1997. Analysis of the human TATA binding protein promoter and identification of an ets site critical for activity. *Nucleic Acids Res.* 25：2485-2494.

Freiman R. N. and Tjian R. 2003. Regulating the regulators：Lysine modifications make their mark. *Cell* 112：11-17.

Frontini M.，Soutoglou E.，Argentini M.，Bole-Feysot C.，Jost B.，Scheer E.，and Toes L. 2005. TAF9b（formerly TAF9L）is a bona fide TAF that has unique and overlapping roles with TAF9. *Mol. Cell. Biol.* 25：4638-4649.

Fryer C. J.，Lamar E.，Turbachova I.，Kintner C.，and Jones K. A. 2002. Mastermind mediates chromatin-spe-

cific transcription and turnover of the Notch enhancer complex. *Genes Dev.* 16: 1397-1411.

Fyodorov D. V. and Kadonaga J. T. 2002. Dynamics of ATP-dependent chromatin assembly by ACF. *Nature* 418: 897-900.

Fyodorov DV. and Kadonaga J. T. 2003. Chromatin assembly in vitro with purified recombinant ACF and NAP-1. *Methods Enzymol.* 371: 499-515.

Gamble M. J., Erdjument-Bromage H., Tempst P., Freedman L. P., and Fisher R. P. 2005. The histone chaperone TAF-1/SET/INHAT is required for transcription in vitro of chromatin templates. *Mol. Cell. Biol.* 25: 797-807.

Gangloff Y. G., Pointud J. C., Thuault S., Caere L., Romier C., Muratoglu S., Brand M., Toes L., Couderc J. L., and Davidson I. 2001. The TFIID components human $TAF_{II}140$ and *Drosophila* BIP2 ($TAF_{II}155$) are novel metazoan homologues of yeast $TAF_{II}47$ containing a histone fold and a PHD finger. *Mol. Cell. Biol.* 21: 5109-5121.

Garrett-Engele G. M., Siegal M. L., Manoli D. S., Williams B. C., Li H., and Baker B. S. 2002. *intersex*, a gene required for female sexual development in *Drosophila*, is expressed in both sexes and functions together with *doublesex* to regulate terminal differentiation. *Development* 129: 4661-4675.

Ge H. and Roeder R. G. 1994. Purification, cloning, and characterizatson of a human coactivator, PC4, that mediates transcriptional activation of class II genes. *Cell* 78: 513-523.

Ge H., Martinez E., Chiang C. M., and Roeder R. G. 1996. Activator-dependent transcription by mammalian RNA polymerase II: In vitro reconstitution with general transcription factors and cofactors. *Methods Enzymol.* 274: 57-71.

Gegonne A., Weissman J. D., Zhou M., Brady J. N., and Singer D. S. 2006. TAF7: A possible transcription initiation check-point regulator. *Proc. Natl. Acad. Sci.* 103: 602-607.

Geles K. G., Freiman R. N., Liu W. L., Zheng S., Voronina E., and Tjian R. 2006. Cell-type-selective induction of *c-jun* by TAF4b directs ovarian-specific transcription networks. *Proc. Natl. Acad. Sci.* 103: 2594-2599.

Georgieva S., Kirschner D. B., Jagla T., Nabirochkina E., Hanke S., Schenkel H., de Lorenzo C., Sinha P., Jagla K., Mechler B., and bra L. 2000. Two novel *Drosophila* TAF_{II}s have homology with human $TAF_{II}30$ and are differentially regulated during development. *Mol. Cell. Biol.* 20: 1639-1648.

Gervais V., Busso D., Wasielewski E., Poterszman A., Egly J. M., Thierry J. C., and Kieffer B. 2001. Solution structure of the N-terminal domain of the human TFIIH MAT1 subunit: New insights into the RING finger family. *J. Biol. Chem.* 276: 7457-7464.

Gervais V., Lamour V., Jawhari A., Frindel F., Wasielewski E., Dubaele S., Egly J. M., Thierry J. C., Kieffer B., and Poterszman A. 2004. TFIIH contains a PH domain involved in DNA nucleotide excision repair. Nat. *Struct. Mol. Biol.* 11: 616-622.

Ghazy M. A., Brodie S. A., Ammerman M. L., Ziegler L. M., and Ponticelli A. S. 2004. Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. *Mol. Cell. Biol.* 24: 10975-10985.

Giardina C. and Lis J. T. 1993. DNA melting on yeast RNA polymeraseII promoters. *Science* 261: 759-762.

Gileadi O., Feaver W. J., and Kornberg R. D. 1992. Cloning of a subunit of yeast RNA polymerase II transcription factor β and CTD kinase. *Science* 257: 1389-1392.

Gim B. S., Park J. M., Yoon J. H., Kang C., and Kim Y. J. 2001. *Drosophila* Med6 is required for elevated expression of a large but distinct set of developmentally regulated genes. *Mol. Cell. Biol.* 21: 5242-5255.

Gnatt AL., Cramer P., Fu J., Bushnell D. A., and Kornberg R. D. 2001. Structural basis of transcription: An RNA polymerase II elongalion complex at 3. 3 Å resolution. *Science* 292: 1876-1882.

Goodrich J. A. and Tjian R. 1994. Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* 77: 145-156.

Gorski K., Carneiro M., and Schibler U. 1986. Tissue-specific in vitro transcription from the mouse albumin promoter. *Cell* 47: 767-776.

Grob P., Cruse M. J., Inouye C., Pens M., Penczek P. A., Tjian R., and Nogales E. 2006. Cryo-electron microscopy studies of human TFIID: Conformational breathing in the integration of gene regulatory cues. *Structure* 14: 511-520.

Gromoller A. and Lehming N. 2000. Srb7p is a physical and physiological target of Tupip. *EMBO J*. 19: 6845-6852.

Gu W., Malik S., Ito M., Yuan C. X., Fondell J. D., Zhang X., Martinez E., Qin J., and Roeder R. G. 1999. A novel human SRB/MED-contaming cofactor complex, SMCC, involved in transcription regulation. *Mol. Cell* 3: 97-108.

Guermah M., Tao Y., and Roeder R. G. 2001. Positive and negative TAF_{II} functions that suggest a dynamic TFIID structure and elicit synergy with traps in activator-induced transcription. *Mol. Cell. Biol.* 21: 6882-6894.

Guermah M., Ge K., Chiang C. M., and Roeder R. G. 2003. The TBN protein, which is essential for early embryonic mouse development, is an inducible TAFII implicated in adipogenesis. *Mol. Cell* 12: 991-1001.

Guermah M., Palhan V. B., Jackett A. J., Chait B. T., and Roeder R. G. 2006. Synergistic functions of SII and p300 in productive activator-dependent transcription of chromatin templates. *Cell* 125: 275-286.

Ha I., Lane W. S., and Reinberg D. 1991. Cloning of a human gene encoding the general transcription initiation factor IIB. *Nature* 352: 689-695.

Hahn S. 1998. The role of TAFs in RNA polymerase II transcription. *Cell* 95: 579-582.

Hahn S. 2004. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* 11: 394-403.

Hahn S., Buratowski S., Sharp P. A., and Guarente L. 1989. Isolation of the gene encoding the yeast TATA binding protein TFIID: A gene identical to the SPT15 suppressor of Ty element insertions. *Cell* 58: 1173-1181.

Hallberg M., Polozkov G. V., Hu G. Z., Beve J., Gustafsson CM., Ronne H., and Bjorklund S. 2004. Site-specific Srb 10-dependent phosphorylation of the yeast Mediator subunit Med2 regulates gene expression from the 2-Jim plasmid. *Proc. NatL Acad. Sci.* 101: 3370-3375.

Hamard P. J., Dalbies-Tran R., Hauss C., Davidson I., Kedinger C., and Chatton B. 2005. A functional interaction between ATF7 and TAF12 that is modulated by TAF4. *Oncogene* 24: 3472-3483.

Hampsey M. 1998. Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol. Mol. Biol. Rev.* 62: 465-503.

Han S. J., Lee J. S., Kang J. S., and Kim Y. J. 2001. Med9/Cse2 and Galil modules are required for transcriptional repression of distinct group of genes. *J. Biol. Chem.* 276: 37020-37026.

Han S. J., Lee Y. C., Gim B. S., Ryu G. H., Park S. J., Lane W. S., and Kim Y. J. 1999. Activator-specific requirement of yeast mediator proteins for RNA polymerase II transcriptional activation. *Mol. Cell. Biol.* 19: 979-988.

Hsrlow E. and Lane D. 1999. *Using antibodies: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Hasegawa J., Endou M., Narita T., Yamada I., Yamaguchi Y., Wada T., and Hands H. 2003. A rapid purification method for human RNA polymerase II by two-step affinity chromatography. I. *Biochem.* 133: 133-138.

Hassan A. H., Neely K. E., and Workman J. L. 2001. Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. *Cell* 104: 817-827.

Hawley D. K. and Roeder R. G. 1985. Separation and partial characterization of three functional steps in transcription initiation by human RNA polymerase II. *J. Biol. Chem.* 260: 8163-8172.

Hawley D. K. and Roeder R. G. 1987. Functional steps in transcription initiation and reinitiation from the major late promoter in a HeLa nuclear extract. *J. Biol. Chem.* 262: 3452-3461.

Hengartner C. J., Myer V. E., Liao S. M., Wilson C. T., Koh S. S., and Young R. A. 1998. Temporal regu-

lation of RNA polymerase II by SrblO and Kin28 cyclmn-dependent kinases. *Mol. Cell* 2：43-53.

Henry N. L.，Sayre M. H.，and Kornberg R. D. 1992. Purification and characterization of yeast RNA polymerase II general initiation factor g. *J. Biol. Chem.* 267：23388-23392.

Hernandez-Hernandez A. and Ferrus A. 2001. Prodos is a conserved transcriptional regulator that interacts with $dTAF_{II}16$ in *Drosophila melanogaster*. *Mol. Cell. Biol.* 21：614-623.

Hieb A. R.，Baran S.，Goodrich J. A.，and Kugel J. F. 2006. An 8 nt RNA triggers a rate-limiting shift of RNA polymerase II complexes into elongation. *EMBO J*. 25：3100-3109.

Hiller M. A.，Lin T. V.，Wood C.，and Fuller MT. 2001. Developmental regulation of transcription by a tissue-specific TAF homolog. *Genes Dev.* 15：1021-1030.

Holstege F. C.，Fiedler U.，and Timmers H. T. 1997. Three transitions in the RNA polymerase II transcription complex during initiation. *EMBOJ*. 16：7468-7480.

Hoistege F. C.，Tantin D.，Carey M.，van der Vliet P. C.，and Timmers H. T. 1995. The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA. *EMBO J*. 14：810-819.

Horikoshi M.，Fujita H.，Wang J.，Takada R.，and Roeder R. G. 1991. Nucleotide and amino acid sequence of RAP30. *Nucleic Acids Res*. 19：5436.

Horikoshi M.，Hai T.，Lin Y. S.，Green M. R.，and Roeder R. G. 1988. Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex. *Cell* 54：1033-1042.

Humbert S.，van Vuuren H.，Lutz Y.，Hoeijmakers J. H.，Egly J. M.，and Moncollin V. 1994. p44 and p34 subunits of the BTF2/TFIIH transcription factor have homologies with SSL1，a yeast protein involved in DNA repair. *EMBO J*. 13：2393-2398.

Inostroza J.，Flores O.，and Reinberg D. 1991. Factors involved in specific transcription by mammalian RNA polymerase II. Purification and functional analysis of general transcription factor TIE. *J. Biol. Chem.* 266：9304-9308.

Irvin J. D. and Pugh B. F. 2006. Genome-wide transcriptional dependence on TAF1 functional domains. *J. Biol. Chem.* 281：6404-6412.

Ito M. and Roeder R. G. 2001. The TRAP/SMCC/Mediator complex and thyroid hormone receptor function. *Trends Endocrinol. Metab.* 12：127-134.

Ito M.，Okano H. J.，Darnell R. B.，and Roeder R. G. 2002. The TRAP 100 component of the TRAP/Mediator complex is essential in broad transcriptional events and development. *EMBO J*. 21：3464-3475.

Ito M.，Yuan C. X.，Okano H. J.，Darnell R. B.，and Roeder R. G. 2000. Involvement of the TRAP220 component of the TRAP/SMCC coactivator complex in embryonic development and thyroid hormone action. *Mol. Cell* 5：683-693.

Ito M.，Yuan C. X.，Malik S.，Gu W.，Fondell J. D.，Yamamura S.，Pu Z. Y.，Zhang X.，Qin J.，and Roeder R. G. 1999. Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators. *Mol. Cell* 3：361-370.

Jacq X.，Brou C.，Lutz Y.，Davidson I.，Chambon P.，and Tora L. 1994. Human TAFII30 is present in a distinct TFIID complex and is required for transcriptional activation by the estrogen receptor. *Cell* 79：107-117.

Janody F.，Martirosyan Z.，Benlali A.，and Treisman J. E. 2003. Two subunits of the *Drosophila* mediator complex act together to control cell affinity. *Development* 130：3691-3701.

Jawhari A.．Lame J. P.，Dubaele S.，Lamour V.，Poterszman A.，Coin F.，Moras D.，and Egly J. M. 2002a. p52 Mediates XPB function within the transcription/repair factor TFIIH. *J. Biol. Chem.* 277：31761-31767.

Jawhari A.，Uhring M.，Crucifix C.，Fribourg S.，Schultz P.，Poterszman A.，Egly J. M.，and Moras D. 2002b. Expression of FLAG fusion proteins in insect cells：Application to the multi-subunit transcription/DNA repair factor TFIIH. *Protein Expr. Purif*. 24：513-523.

Jawhari A.，Uhring M.，De Carlo S.，Crucifix C.，Tocchini-Valentini G.，Moras D.，Schultz P.，and Potersz-

man A. 2006. Structure and oligomeric state of human transcription factor TFIIE. *EMBO Rep*. 7: 500-505.

Jiang Y. and Gralla J. D. 1993. Uncoupling of initiation and reinitiation rates during HeLa RNA polymerase II transcription in vitro. *Mol. Cell. Biol.* 13: 4572-4577.

Jiang Y., Triezenberg S. J., and Gralla J. D. 1994. Defective transcriptional activation by diverse VP16 mutants associated with a common inability to form open promoter complexes. *J. Biol. Chem.* 269: 5505-5508.

Jiang Y. W., Veschambre P., Erdjument-Bromage H., Tempst P., Conaway J. W., Conaway R. C., and Kornberg RD. 1998. Mammalian mediator of transcriptional regulation and its possible role as an end-point of signal transduction pathways. *Proc. Natl. Acad. Sci.* 95: 8538-8543.

Johnson F. B. and Krasnow M. A. 1992. Differential regulation of transcription preinitiation complex assembly by activator and repressor homeo domain proteins. Genes *Dev.* 6: 2177-2189.

Johnson K. M. and Carey M. 2003. Assembly of a mediator/TFIID/ TFITA complex bypasses the need for an activator. *Curr. Biol.* 13: 772-777.

Johnson K. M., Wang J., Smailwood A., and Carey M. 2004. The immobilized template assay for measuring cooperativity in eukaryotic transcription complex assembly. *Methods Enzymol*. 380: 207-219.

Johnson K. M., Wang J., Smallwood A., Arayata C., and Carey M. 2002. TFIID and human mediator coactivator complexes assemble cooperatively on promoter DNA. *Genes Dev.* 16: 1852-1863.

Jona G., Livi L. L., and Gileadi O. 2002. Mutations in the RING domain of TFB3, a subunit of yeast transcription factor IIH, reveal a role in cell cycle progression. *J. Biol. Chem.* 277: 39409-39416.

Jones K. A. 1997. Taking a new TAK on tat transactivation. *Genes Dev.* 11: 2593-2599.

Kamakaka R. T. and Kadonaga J. T. 1994. The soluble nuclear fraction, a highly efficient transcription extract from *Drosophila* embryos. *Methods Cell Biol*. 44: 225-235.

Kang J. S., Kim S. H., Hwang M. S., Han S. J., Lee Y. C., and Kim Y. J. 2001. The structural and functional organization of the yeast mediator complex. *J. Biol. Chem.* 276: 42003-42010.

Kasher M. S., Pintel O., and Ward D. C. 1986. Rapid enrichment of HeLa transcription factors IIIB and IIIC by using affinity chromatography based on avidin-biotin interactions. *Mol. Cell. Biol.* 6: 3117-3127.

Kassavetis G. A., Blanco J. A., Johnson T. E., and Geiduschek E. P. 1992. Formation of open and elongating transcription complexes by RNA polymerase III. *J. Mol. Biol.* 226: 47-58.

Kelleher R. J., 3rd, Flanagan P. M., and Kornberg R. D. 1990. A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* 61: 1209-1215.

Keogh M. C., Cho E. J., Podolny V., and Buratowski S. 2002. Kin28 is found within TFIIH and a Kin28-Ccll-Tfls3 trimer complex with differential sensitivities to T-loop phosphorylation. *Mol. Cell. Biol.* 22: 1288-1297.

Kershnar E., Wu S. Y., and Chiang C. M. 1998. Immunoaffinity purification and functional characterization of human transcription factor IIH and RNA polymerase II from clonal cell lines that conditionally express epitope-tagged subunits of the multiprotein complexes. *J. Biol. Chem.* 273: 34444-34453.

Killeen M. T. and Greenblatt J. F. 1992. The general transcription factor RAP30 binds to RNA polymerase II and prevents it from binding nonspecifically to DNA. *Mol. Cell. Biol.* 12: 30-37.

Kim S., Xu X., Hecht A., and Bayer T. G. 2006. Mediator is a transducer of Wnt/beta-catenin signaling. *J. Biol. Chem.* 281: 14066-14075.

Kim T. K. and Maniatis T. 1997. The mechanism of transcriptional synergy of an in vitro assembled interferon-β enhanceosome. *Mol. Cell* 1: 119-129.

Kim T. K., Ebright R. H., and Reinberg D. 2000. Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* 288: 1418-1422.

Kim T. K., Lagrange T., Wang Y. H., Griffith J. D., Reinberg O., and Ebright R. H. 1997. Trajectory of DNA in the RNA polymerase II transcription preinitiation complex. *Proc. Natl. Acad. Sci.* 94: 12268-12273.

Kim T. W., Kwon Y. I., Kim J. M., Song Y. H., Kim S. N., and Kim Y. J. 2004. MED16 and MED23 of

Mediator are coactivators of lipopolysaccharide-and heat-shock-induced transcriptional activators. *Proc. Natl. Acad. Sci.* 101: 12153-12158.

Kim Y. J., Bjdrklund S., Li Y., Sayre M. H., and Kornberg R. D. 1994. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* 77: 599-608.

Kimura M. and Ishihama A. 2004. Tfg3, a subunit of the general transcription factor TFIIF in *Schizosaccharomyces pombe*, functions under stress conditions. *Nucleic Acids Res.* 32: 6706-6715.

Kimura M., Suzuki H., and Ishihama A. 2002. Formation of a carboxy-terminal domain phosphatase (Fcp1) /TFIIF/RNA polymerase II (pol II) complex in *Schizosaccharomyces ponibe* involves direct interaction between Fcpl and the Rpb4 subunit of pol II. *Mol. Cell. Biol.* 22: 1577-1588.

Kobor M. S. and Greenblatt J. 2002. Regulation of transcription elongation by phosphorylation. *Biochim. Biophys. Acta* 1577: 261-275.

Koh S. S., Ansari A. Z., Ptashne M., and Young R. A. 1998. An activator target in the RNA polymerase II holoenzyme. *Mol. Cell* 1: 895-904.

Korsisaari N., Rossi D. J., Paetau A., Charnay P., Henkemeyer M., and Makela T. P. 2002. Conditional ablation of the Mati subunit of TFIIH in Schwann cells provides evidence that Mati is not required for general transcription. *J. Cell Sci.* 115: 4275-4284.

Kraemer S. M., Ranallo R. T., Ogg R. C., and Stargell L. A. 2001. TFIIA interacts with TFIID via association with TATA-binding protein and TAF40. *Mol. Cell. Biol.* 21: 1737-1746.

Kraus W. L. and Kadonaga J. T. 1998. p300 and estrogen receptor cooperatively activate transcription via differential enhancement of initiation and reinitiation. *Genes Dev.* 12: 331-342.

Kumar K. P., Akoulitchev S., and Reinberg D. 1998. Promoter-proximal stalling results from the inability to recruit transcription factor IIH to the transcription complex and is a regulated event. *Proc. Natl. Acad. Sci.* 95: 9767-9772.

Kwon J. Y., Park J. M., Gim B. S., Han S. J., Lee J., and Kim Y. J. 1999. *Caenorhabditis elegans* mediator complexes are required for developmental-specific transcriptional activation. *Proc. Natl. Acad. Sci.* 96: 14990-14995.

Lagrange T., Kapanidis A. N., Tang H., Reinberg D., and Ebright R. H. 1998. New core promoter element in RNA polymerase II-dependent transcription: Sequence-specific DNA binding by transcription factor IIB. *Genes Dev.* 12: 34-44.

Lagrange T., Kim T. K., Orphanides G., Ebright Y. W., Ebright R. H.,. and Reinberg D. 1996. High-resolution mapping of nucleoprotein complexes by site-specific protein-DNA photocrosslinking: Organization of the human TBP-TFIIA-TFIIB-DNA quaternary complex. *Proc. Natl. Acad. Sci.* 93: 10620-10625.

Laine J. P. and Egly J. M. 2006. Initiation of DNA repair mediated by a stalled RNA polymerase IIO. *EMBO J.* 25: 387-397.

Larivière L., Geiger S., Hoeppner S., Röther S., Sträer K., and Cramer P. 2006. Structure and TBP binding of the Mediator head subcomplex Med8-Med18-Med20. *Nat. Struct. Mol. Biol.* 13: 895-901

Larochelle S., Pandur I., Fisher R. P., Salz H. K., and Suter B. 1998. Cdk7 is essential for mitosis and for in vivo Cdk-activating kinase activity. *Genes Dev.* 12: 370-381.

Lau J. F., Nusinzon I., Burakov D., Freedman L. P., and Horvath C. M. 2003. Role of metazoan mediator proteins in interferon-responsive transcription. *Mol. Cell. Biol.* 23: 620-628.

Lavigne M., Francis N. J., King I. F., and Kingston R. E. 2004. Propagation of silencing; Recruitment and repression of naive chromatin in trans by polycomb repressed chromatin. *Mol. Cell* 13: 415-425.

Lee K. A., Bindereif A., and Green M. R. 1988. A small-scale procedure for preparation of nuclear extracts that support efficient transcription and pre-mRNA splicing. *Gene Anal. Tech.* 5: 22-31.

Lee K. M., Miklos I., Du H., Watt S., Szilagyi Z., Saiz I. E., Madabhushi R., Penkett C. J., Sipiczki M.,

Bahler J., and Fisher R. P. 2005. Impairment of the TFIIH-associated CDK-activating kinase selectively affects cell cycle-regulated gene expression in fission yeast. *Mol. Biol. Cell* 16: 2734-2745.

Lee T. I., Wyrick J. J., Koh S. S., Jennings E. G., Gadbois E. L., and Young R. A. 1998. Interplay of positive and negative regulators in transcription initiation by RNA polymerase II holoenzyme. *Mol. Cell. Biol.* 18: 4455-4462.

Lee T. I., Causton H. C., Holstege F. C., Shen W. C., Hannett N., Jennings E. G., Winston F., Green MR., and Young R. A. 2000. Redundant roles for the TFIID and SAGA complexes in global transcription. *Nature* 405: 701-704.

Lee Y. C. and Kim Y. J. 1998. Requirement for a functional interaction between mediator components Med6 and Srb4 in RNA polymerase II transcription. *Mol. Cell. Biol.* 18: 5364-5370.

Lemaire M. and Collart M. A. 2000. The TATA-binding protein-associated factor yTafII 19p functionally interacts with components of the global transcriptional regulator Ccr4-Not complex and physically interacts with the Not5 subunit. *J. Biol. Chem.* 275: 26925-26934.

Levedakou E. N., He M., Baptist E. W., Craven R. J., Cance W. G., Welcsh P. L., Simmons A., Naylor S. L., Leach R. J., Lewis T. B., et al. 1994. Two novel human serine/threonine kinases with homologies to the cell cycle regulating *Xenopus* MO15, and NIMA kinases: Cloning and characterization of their expression pattern. *Oncogene* 9: 1977-1988.

Lewis B. A. and Reinberg D. 2003. The mediator coactivator complex: Functional and physical roles in transcriptional regulation. *J. Cell Sci.* 116: 3667-3675.

Li Y. J., Fu X. H., Liu D. P., and Liang C. C. 2004. Opening the chromatin for transcription. *Int. J. Biochem. Cell Biol.* 36: 1411-1423.

Lieberman P. 1994. Identification of functional targets of the Zta transcriptional activator by formation of stable preinitiation complex intermediates. *Mol. Cell. Biol.* 14: 8365-8375.

Lieberman P. M. and Berk A. J. 1994. A mechanism for TAFs in transcriptional activation: Activation domain enhancement of TFIID-TFIIA-promoter DNA complex formation. *Genes Dev.* 8: 995-1006.

Lin Y. C. and Gralla J. D. 2005. Stimulation of the XPB ATP-dependent helicase by the beta subunit of TFIIE. *Nucleic Acids Res.* 33: 3072-3081.

Lin Y. C., Choi W. S., and Gralla J. D. 2005. TFIIH XPB mutants suggest a unified bacterial-like mechanism for promoter opening but not escape. *Nat. Struct. Mol. Biol.* 12: 603-607.

Liu Y., Kung C., Fishb urn J., Ansari A. Z., Shokat K. M., and Hahn S. 2004. Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. Mol. *Cell. Biol.* 24: 1721-1735.

Liu Y., Ando S., Xia X., Yao R., Kim M., Fondell J., and Yen P. M. 2005. p62, A TFIIH subunit, directly interacts with thyroid hormone receptor and enhances T3-mediated transcription. *Mol. Endocrinol.* 19: 879-884.

Lorch Y., Beve J., Gustafason C. M., Myers L. C., and Kornberg R. D. 2000. Mediator-nucleosome interaction. *Mol. Cell* 6: 197-201.

Loyola A. and Reinberg D. 2003. Histone deposition and chromatin assembly by RSF. *Methods* 31: 96-103.

Loyola A., LeRoy G., Wang Y. H., and Reinberg D. 2001. Reconstitution of recombinant chromatin establishes a requirement for histone-tail modifications during chromatin assembly and transcription. *Genes Dev.* 15: 2837-2851.

Lu Z., Ansari AZ., Lu X., Ogirala A., and Ptashne M. 2002. A target essential for the activity of a nonacidic yeast transcriptional activator. *Proc. Natl. Acad. Sci.* 99: 8591-8596.

Luse D. S. and Jacob G. A. 1987. Abortive initiation by RNA polymerase II in vitro at the adenovirus 2 major late promoter. *J. Biol. Chem.* 262: 14990-14997.

Lusser A., Urwin D. L., and Kadonaga J. T. 2005. Distinct activities of CHD1 and ACF in ATP-dependent chro-

matin assembly. *Nat. Struct. Mol. Biol.* 12: 160-166.

Ma D., Watanabe H., Mermelstein F., Admon A., Oguri K., Sun X., Wads T., Imai T., Shiroya T., Reinberg D., et al. 1993. Isolation of a cDNA encoding the largest subunit of TFIIA reveals functions important for activated transcription. *Genes Dev.* 7: 2246-2257.

Maile T., Kwoczynski S., Katzenberger R. J., Wassarman D. A., and Sauer F. 2004. TAF1 activates transcription by phosphorylation of serine 33 in histone H2B. *Science* 304: 1010-1014.

Makela T. P., Tassan J. P., Nigg E. A., Frutiger S., Hughes G. J., and Weinberg R. A. 1994. A cyclin associated with the CDK-activating kinase MO15. *Nature* 371: 254-257.

Maldonado E., Drapkin R., and Reinberg D. 1996. Purification of human RNA polymerase II and general transcription factors. *Methods Enzymol.* 274: 72-100.

Malik S. and Roeder R. G. 2000. Transcriptionai regulation through Mediator-like coactivators in yeast and metazoan cells. *Trends Biochem. Sci.* 25: 277-283.

Malik S. and Roeder R. G. 2003. Isolation and functional characterization of the TRAP/mediator complex. *Methods Enzymol.* 364: 257-284.

Malik S. and Roeder R. G. 2005. Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem. Sci.* 30: 256-263.

Malik S., Baek H. J., Wu W., and Roeder R. G. 2005. Structural and functional characterization of PC2 and RNA polymerase II-associated subpopulations of metazoan Mediator. *MoL Cell. Biol.* 25: 2117-2129.

Manley J. L., Fire A., Samuels M., and Sharp P. A. 1983. In vitro transcription: Whole-cell extract. *Methods Enzymol.* 101: 568-582.

Manley J. L., Fire A., Cano A., Sharp P. A., and Gefter M. L. 1980. DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. *Proc. Natl. Acad. Sci.* 77: 3855-3859.

Margueron R., Trojer P., and Reinberg D. 2005. The key to development: Interpreting the histone code? *Curr. Opin. Genet. Dev.* 15: 163-176.

Marinoni J. C., Roy R., Vermeulen W., Miniou P., Lutz Y., Weeda G., Seroz T., Gomez D. M., Hoeijmakers J. H., and Egly J. M. 1997. Cloning and characterization of p52, the fifth subunit of the core of the transcription/DNA repair factor TFIIH. *EMBO J.* 16: 1093-1102.

Martin J., Halenbeck R., and Kaufmann J. 1999. Human transcription factor $hTAF_{II}150$ (CIF150) is involved in transcriptional regulation of cell cycle progression. *Mol. Cell. Biol.* 19: 5548-5556.

Martinez E. 2002. Multi-protein complexes in eukaryotic gene transcription. *Plant Mol. Biol.* 50: 925-947.

Matangkasombut O., Auty R., and Buratowski S. 2004. Structure and function of the TFIID complex. *Adv. Protein Chem.* 67: 67-92.

Matsui T., Segall J., Well P. A., and Roeder R. G. 1980. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J. Biol. Chem.* 255: 11992-11996.

Maxon ME., Goodrich J. A., and Tjian R. 1994. Transcription factor IIE binds preferentially to RNA polymerase ha and recruits TFIIH: A model for promoter clearance. *Genes Dev.* 8: 515-524.

May M., Mengus G., Lavigne A. C., Chambon P., and Davidson I. 1996. Human TAF (II28) promotes transcriptional stimulation by activation function 2 of the retinoid X receptors. *EMBO J.* 15: 3093-3104.

Mayall T. P., Sheridan P. L., Montminy M. R., and Jones K. A. 1997. Distinct roles for P-CREB and LEF-1 in $TCR\alpha$ enhancer assembly and activation on chromatin templates in vitro. *Genes Dev.* 11: 887-899.

Meinhart A., Blobel J., and Cramer P. 2003. An extended winged helix domain in general transcription factor E/$IIE\alpha$. *J. Biol. Chem.* 278: 48267-48274.

Mellor J. 2005. The dynamics of chromatin remodeling at promoters. *Mol. Cell* 19: 147-157.

Mengus G., May M., Carre L., Chambon P., and Davidson I. 1997. Human $TAF_{II}35$ potentiates transcriptional activation by the AF-2s of the retinoic acid, vitamin D3, and thyroid hormone receptors in mammalian cells.

Genes Dev. 11: 1381-1395.

Mengus G., May M., Jacq X., Staub A., Tora L., Chambon P., and Davidson I. 1995. Cloning and characterization of $hTAF_{II}18$, $hTAF_{II}20$ and $hTAF_{II}28$: Three subunits of the human transcription factor TFIID. *EMBO J.* 14: 1520-1531.

Metzger D., Scheer E., Soldatov A., and Tora L. 1999. Mammalian $TAF_{II}30$ is required for cell cycle progression and specific cellular differentiation programmes. *EMBO J.* 18: 4823-4834.

Mitsiou D. J. and Stunnenberg H. G. 2000. TAC, a TBP-sans-TAFs complex containing the unprocessed TFIIA $\alpha\beta$ precursor and the TFIIAγ subunit. *Mol. Cell* 6: 527-537.

Mittler G., Stuhler T., Santolin L., Uhlmann T., Kremmer E., Lottspeich F., Berti L., and Meisterernst M. 2003. A novel docking site on Mediator is critical for activation by VP16 in mammalian cells. *EMBO J.* 22: 6494-6504.

Mo X., Kowenz-Leutz E., Xu H., and Leutz A. 2004. Ras induces mediator complex exchange on C/EBPβ. *Mol. Cell* 13: 241-250.

Mohan W. S., Jr., Scheer E., Wendling O., Metzger D., and Tora L. 2003. TAF10 ($TAF_{II}30$) is necessary for TFIID stability and early embryogenesis in mice. *Mol. Cell. Biol.* 23: 4307-4318.

Moreland R. J., Tirode F., Yan Q., Conaway J. W., Egly J. M., and Conaway R. C. 1999. A role for the TFIIH XPB DNA helicase in promoter escape by RNA polymerase II. *J. Biol. Chem.* 274: 22127-22130.

Mulholland N. M., King I. F., and Kingston R. E. 2003. Regulation of Polycomb group complexes by the sequence-specific DNA binding proteins Zeste and GAGA. *Genes Dev.* 17: 2741-2746.

Muller F. and Tora L. 2004. The multicoloured world of promoter recognition complexes. *EMBO J.* 23: 2-8.

Muncke N., Jung C., Rudiger H., Ulmer H., Roeth R., Hubert A., Goldmuntz E., Driscoll D., Goodship J., Schon K., and Rappold G. 2003. Missense mutations and gene interruption in *PROSIT*240, a novel *TRAP240-like* gene, in patients with congenital heart defect (transposition of the great arteries). *Circulation* 108: 2843-2850.

Munz C., Psichari E., Mandilis D., Lavigne A. C., Spiliotaki M., Oehler T., Davidson I., Tora L., Angel P., and Pintzas A. 2003. TAF7 ($TAF_{II}55$) plays a role in the transcription activation by c-Jun. *J. Biol. Chem.* 278: 21510-21516.

Musante L., Bartsch O., Ropers H. H., and Kalscheuer V. M. 2004. cDNA cloning and characterization of the human *THRAP*2 gene which maps to chromosome 12q24, and its mouse ortholog *Thrap*2. *Gene* 332: 119-127.

Myer V. E. and Young R. A. 1998. RNA polymerase II holoenzymes and subcomplexes. *J. Biol. Chem.* 273: 27757-27760.

Myers L. C., Gustafsson C. M., Hayashibara K. C., Brown P. O., and Kornberg R. D. 1999. Mediator protein mutations that selectively abolish activated transcription. *Proc. Natl. Acad. Sci.* 96: 67-72.

Naar A. M., Taatjes D. J., Zhai W., Nogales E., and Tjian R. 2002. Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. *Genes Dev.* 16: 1339-1344.

Naar A. M., Beaurang PA., Zhou S., Abraham S., Solomon W., and Tjian R. 1999. Composite co-activator ARC mediates chromatindirected transcriptional activation. *Nature* 398: 828-832.

Nakajima N., Horikoshi M., and Roeder R. G. 1988. Factors involved in specific transcription by mammalian RNA polymerase II: Purification, genetic specificity, and TATA box-promoter interactions of TFIID. *Mol. Cell. Biol.* 8: 4028-4040.

Naumovski L. and Friedberg E. C. 1982Molecular cloning of eucaryotic genes required for excision repair of UV-irradiated DNA: Isolation and partial characterization of the *RAD*3 gene of *Saccharornyces cerevisiae*. *J. Bacteriol.* 152: 323-331.

Nikolov D. B., Chen H., Halay E. D., Usheva A. A., Hisatake K., Lee D. K., Roeder R. G., and Burley S. K. 1995. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* 377: 119-128.

Oelgeschlager T., Chiang C. M., and Roeder R. G. 1996. Topology and reorganization of a human TFIID-promoter complex. *Nature* 382: 735-738.

Oelgeschlager T., Tao Y., Kang Y. K., and Roeder R. G. 1998. Transcription activation via enhanced preinitiation complex assembly in a human cell-free system lacking TAFIIs. *Mol. Cell* 1: 925-931.

Ogata K., Sato K., and Tahirov T. H. 2003. Eukaryotic transcriptional regulatory complexes: Cooperativity from near and afar. *Curr. Opin. Struct. Biol.* 13: 40-48.

Ogryzko V. V., Kotani T., Zhang X., Schiltz R. L., Howard T., Yang X. J., Howard B. H., Qin J., and Nakatani Y. 1998. Histone-like TAFs within the PCAF histone acetylase complex. *Cell* 94: 35-44.

Ohkuma Y. and Roeder R. G. 1994. Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* 368: 160-163.

Ohkuma Y., Sumimoto H., Horikoshi M., and Roeder R. G. 1990. Factors involved in specific transcription by mammalian RNA polymerase II Purification and characterization of general transcription factor TFIIE. *Proc. Natl. Acad. Sci.* 87: 9163-9167.

Ohkuma Y., Sumimoto H., Hoffmann A., Shimasaki S., Horikoshi M., and Roeder R. G. 1991. Structural motifs and potential sigma homologies in the large subunit of human general transcription factor TFIIE. *Nature* 354: 398-401.

Okamoto T., Yamamoto S., Watanabe Y., Ohta T., Hanaoka F., Roeder R. G., and Ohkuma Y. 1998. Analysis of the role of TFIIE in transcriptional regulation through structure-function studies of the TFIIEβ subunit. *J. Biol. Chem.* 273: 19866-19876.

Okuda M., Tanaka A., Arai Y., Satoh M., Okamura H., Nagadoi A., Hanaoka F., Ohkuma Y., and Nishimura Y. 2004. A novel zinc finger structure in the large subunit of human general transcription factor TFIIE. *J. Biol. Chem.* 279: 51395-51403.

Orphanides G. and Reinberg D. 2002. A unified theory of gene expression. *Cell* 108: 439-451.

Orphanides G., Lagrange T., and Reinberg D. 1996. The general transcription factors of RNA polymerase II. *Genes Dev.* 10: 2657-2683.

Osborn L., Kunkel S., and Nabel G. J. 1989. Tumor necrosis factor α and interleukin 1 stimulate the human immunodeficiency virus enhancer by activation of the nuclear factor κB. *Proc. Natl. Acad. Sci.* 86: 2336-2340.

Ozer J., Moore P. A., and Lieberman P. M. 2000. A testis-specific transcription factor IIA (TFIIAτ) stimulates TATA-binding protein-DNA binding and transcription activation. *J. Biol. Chem.* 275: 122-128.

Ozer J., Moore P. A., Bolden A. H., Lee A., Rosen C. A., and Lieberman P. M. 1994. Molecular cloning of the small (γ) subunit of human TFIIA reveals functions critical for activated transcription. *Genes Dev.* 8: 2324-2335.

Pal M., Ponticelli A. S., and Luse D. S. 2005. The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. *Mol. Cell* 19: 101-110.

Pan G. and Greenblatt J. 1994. Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation Site. *J. Biol. Chem.* 269: 30101-30104.

Paoletti A. C., Parmely T. J., Tomomori-Sato C., Sato S., Zhu D., Conaway R. C., Conaway J. W., Florens L., and Washburn M. P. 2006. Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc. Natl. Acad. Sci.* 103: 18928-18933.

Papamichos-Chronakis M., Conlan R. S., Gounalaki N., Copf T., and Tzamarias D. 2000. Hrs1/Med3 is a Cyc8-Tupl corepressor target in the RNA polymerase II holoenzyme. *J. Biol. Chem.* 275: 8397-8403.

Park E., Guzder S. N., Koken M. H., Jaspers-Dekker C, Weeda G., Hoeijmakers J. H., Prakash S., and Prakash L. 1992. RAD25 (SSL2), the yeast homolog of the human xeroderma pigmentosum group B DNA repair gene, is essential for viability. *Proc. Natl. Acad. Sci.* 89: 11416-11420.

Park J. M., Kim H. S., Han S. J., Hwang M. S., Lee Y. C., and Kim Y. J. 2000. In vivo requirement of acti-

vator-specific binding targets of mediator. *Mol. Cell. Biol.* 20: 8709-8719.

Parker C. S. and Topol J. 1984. A *Drosophila* RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. *Cell* 36: 357-369.

Pavri R., Zhu B., Li G., Trojer P., Mandal S., Shilatifard A., and Reinberg D. 2006. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* 125: 703-717.

Pavri R., Lewis B., Kim T. K., Dilworth F. J., Erdjument-Bromage H., Tempst P., de Murcia G., Evans R., Chambon P., and Reinberg D. 2005. PARP-1 determines specificity in a retinoid signaling pathway via direct modulation of mediator. *Mol. Cell* 18: 83-96.

Pazin M. J., Kamakaka R. T., and Kadonaga J. T. 1994. ATP-dependent nucleosome reconfiguration and transcriptional activation from preassembled chromatin templates. *Science* 266: 2007-2011.

Pazin M. J., Sheridan P. L., Cannon K., Cao Z., Keck J. G., Kadonaga J. T., and Jones K. A. 1996. NF-κB-mediated chromatin reconfiguration and transcriptional activation of the HIV-1 enhancer in vitro. *Genes Dev.* 10: 37-49.

Perletti L., Kopf E., Carré L., and Davidson I. 2001. Coordinate regulation of RARγ2, TBP, and $TAF_{II}135$ by targeted proteolysis during retinoic acid-induced differentiation of F9 embryonal carcinoma cells. *BMC Mol. Biol.* 2: 4.

Peterson C. L. and Laniel M. A. 2004. Histones and histone modifications. *Curr. Biol.* 14: R546-551.

Peterson M. G., Inostroza J., Maxon M. E., Flores O., Admon A., Reinberg D., and Tjian R. 1991. Structure and functional properties of human general transcription factor IIE. *Nature* 354: 369-373.

Phelan M. L., Sif S., Narlikar G. J., and Kingston R. E. 1999. Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol. Cell* 3: 247-253.

Pinto I., Ware D. E., and Hampsey M. 1992. The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo. *Cell* 68: 977-988.

Pointud J. C., Mengus G., Brancorsini S., Monaco L., Parvinen M., Sassone-Corsi P., and Davidson I. 2003. The intracellular localisation of TAF7L, a paralogue of transcription factor TFIID subunit TAF7, is developmentally regulated during male germ-cell differentiation. *J. Cell Sci.* 116: 1847-1858.

Price D. H. 2000. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol. Cell. Biol.* 20: 2629-2634.

Rachez C., Lemon B. D., Suldan Z., Bromleigh V., Gamble M., Naar A. M., Erdjument-Bromage H., Tempst P., and Freedman L. P. 1999. Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex. *Nature* 398: 824-828.

Ranish J. A., Lane W. S., and Hahn S. 1992. Isolation of two genes that encode subunits of the yeast transcription factor IIA. *Science* 255: 1127-1129.

Ranish J. A., Yudkovsky N., and Hahn S. 1999. Intermediates in formation and activity of the RNA polymerase II preinitiation complex: Holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev.* 13: 49-63.

Rashi-Elkeles S., Elkon R., Weizman N., Linhart C., Amariglio N., Sternberg G., Rechavi G., Barzilai A., Shamir R., and Shiloh Y. 2006. Parallel induction of ATM-dependent pro-and antiapoptotic signals in response to ionizing radiation in murine lymphoid tissue. *Oncogene* 25: 1584-1592.

Rau M. J., Fischer S., and Neumann C. J. 2006. Zebrafish Trap230/ Med 12 is required as a coactivator for Sox9-dependent neural crest, cartilage and ear development. *Dev. Biol.* 296: 83-93.

Reese J. C. 2003. Basal transcription factors. *Curr. Opin. Genet. Dev.* 13: 114-118.

Reese J. C. and Green M. R. 2001. Genetic analysis of TAF68/61 reveals links to cell cycle regulators. *Yeast* 18: 1197-1205.

Reeves W. M. and Hahn S. 2003. Activator-independent functions of the yeast mediator sin4 complex in preinitiation

complex formation and transcription reinitiation. *Mol. Cell. Biol.* 23: 349-358.

Reid J., Murray I., Watt K., Betney R., and McEwan I. J. 2002. The androgen receptor interacts with multiple regions of the large subunit of general transcription factor TFIIF. *J. Biol. Chem.* 277: 41247-41253.

Reinberg D. and Roeder R. G. 1987. Factors involved in specific transcription by mammalian RNA polymerase II. Purification and functional analysis of initiation factors IIB and IIE. *J. Biol. Chem.* 262: 3310-3321.

Riedl T., Hanaoka F., and Egly J. M. 2003. The comings and goings of nucleotide excision repair factors on damaged DNA. *EMBO J.* 22: 5293-5303.

Roberts S. G., Choy B., Walker S. S., Lin Y. S., and Green M. R. 1995. A role for activator-mediated TFIIB recruitment in diverse aspects of transcriptional regulation. *Curr. Biol.* 5: 508-516.

Roeder R. G. 1996. Nuclear RNA polymerases: Role of general initiation factors and cofactors in eukaryotic transcription. *Methods Enzymol.* 273: 165-171.

Roeder R. G. 2005. Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett.* 579: 909-915.

Roeder R. G. and Rutter W. J. 1969. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* 224: 234-237.

Rossi D. J., Londesborough A., Korsisaari N., Pihlak A., Lehtonen E., Henkemeyer M., and Makela T. P. 2001. Inability to enter S phase and defective RNA polymerase II CTD phosphorylation in mice lacking Mat1. *EMBO J.* 20: 2844-2856.

Ryu S., Zhou S., Ladurner A. G., and Tjian R. 1999. The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. *Nature* 397: 446-450.

Sakurai H. and Fukasawa T. 2000. Functional connections between mediator components and general transcription factors of Saccharomyces cerevisiae. *J. Biol. Chem.* 275: 37251-37256.

Sambrook J. and Russell D. W. 2001. *Molecular cloning: A laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Samuels M., Fire A., and Sharp P. A. 1984. Dinucleotide priming of transcription mediated by RNA polymerase II. *J. Biol. Chem.* 259: 2517-2525.

Sandaltzopoulos R. and Becker P. B. 1998. Heat shock factor increases the reinitiation rate from potentiated chromatin templates. *Mol. Cell. Biol.* 18: 361-367.

Sasse-Dwight S. and Gralla J. D. 1989. $KMnO_4$ as a probe for lac promoter DNA melting and mechanism in vivo. *J. Biol. Chem.* 264: 8074-8081.

Sato S., Tomomori-Sato C., Banks C. A., Parmely T. J., Sorokina I., Brower C. S., Conaway R. C., and Conaway J. W 2003. A mammalian homolog of *Drosophila melanogaster* transcriptional coactivator *intersex* is a subunit of the mammalian Mediator complex. *J. Biol. Chem.* 278: 49671-49674.

Sato S., Tomomori-Sato C., Parmely T. J., Florens L., Zybailov B., Swanson S. K., Banks C. A., Tin J., Cai Y., Washburn M. P., et al. 2004. A set of consensus mammalian mediator subunits identified by multidimensional protein identification technology. *Mol. Cell* 14: 685-691.

Sawadogo M. and Roeder R. G. 1984. Energy requirement for specific transcription initiation by the human RNA polymerase II system. *J. Biol. Chem.* 259: 5321-5326.

Sawadogo M. and Roeder R. G. 1985a. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* 43: 165-175.

Sawadogo M. and Roeder R. G. 1985b. Factors involved in specific transcription by human RNA polymerase II: Analysis by a rapid and quantitative in vitro assay. *Proc. Natl. Acad. Sci.* 82: 4394-4398.

Schewe D. M., Leupold J. H., Boyd D. D., Lengyel E. R., Wang H., Gruetzner K. U., Schildberg F. W., Jauch K. W., and Allgayer H. 2003. Tumor-specific transcription factor binding to an activator protein-2/Sp1 element of the urokinase-type plasminogen activator receptor promoter in a first large series of resected gastroin-

testinal cancers. *Clin. Cancer Res.* 9: 2267-2276.

Schiltz R. L., Mizzen C. A., Vassilev A., Cook R. G., Allis C. D., and Nakatani Y. 1999. Overlapping but distinct patterns of histone acetylation by the human coactivators p300 and PCAF within nucleosomal substrates. *J. Biol. Chem.* 274: 1189-1192.

Schnitzler G., Sif S., and Kingston R. E. 1998. Human SWI/SNF inter-converts a nucleosome between its base state and a stable remodeled state. *Cell* 94: 17-27.

Schultz P., Fribourg S., Poterszman A., Mallouh V., Moras D., and Egly J. M. 2000. Molecular structure of human TFIIH. *Cell* 102: 599-607.

Schwartz B. E., Larochelle S., Suter B., and Lis J. T. 2003. Cdk7 is required for full activation of *Drosophila* heat shock genes and RNA polymerase II phosphorylation in vivo. *Mol. Cell. Biol.* 23: 6876-6886.

Scopes R. K. 1994. *Protein purification: Principles and practice*, 3rd ed. Springer, New York.

Segall J., Matsui T., and Roeder R. G. 1980. Multiple factors are required for the accurate transcription of purified genes by RNA polymerase III. *J. Biol. Chem.* 255: 11986-11991.

Selleck W., Howley R., Fang Q., Podolny V., Fried M. G., Buratowski S., and Tan S. 2001. A histone fold TAF octamer within the yeast TFIID transcriptional coactivator. *Nat. Struct. Biol.* 8: 695-700.

Seroz T., Perez C., Bergmann E., Bradsher J., and Egly J. M. 2000. p44/SSL1, the regulatory subunit of the XPDIRAD3 helicase, plays a crucial role in the transcriptional activity of TFIIH. *J. Biol. Chem.* 275: 33260-33266.

Sheridan P. L., Mayall T. P., Verdin E., and Jones K. A. 1997. Histone acetyltransferases regulate HIV-1 enhancer activity in vitro. *Genes Dev.* 11: 3327-3340.

Sheridan P. L., Sheline C. T., Cannon K., Voz M. L., Pazin M. J., Kadonaga J. T., and Jones K. A. 1995. Activation of the HIV-1 enhancer by the LEF-I HMG protein on nucleosome-assembled DNA in Vitro. *Genes Dev.* 9: 2090-2104.

Shim E. Y., Walker A. K., and Blackwell T. K. 2002a. Broad requirement for the mediator subunit RGR-1 for transcription in the *Caenorhabditis elegans* embryo. *J. Biol. Chem.* 277: 30413-30416.

Shim E. Y., Walker A. K., Shi Y., and Blackwell T. K. 2002b. CDK9/cycin T (P-TEFb) is required in two postinitiation pathways for transcription in the *C. elegans* embryo. *Genes Dev.* 16: 2135-2146.

Silkov A., Wolstein O., Shachar I., and Dikstein R. 2002. Enhanced apoptosis of B and T lymphocytes in TAF_{II} 105 dominant-negative transgenic mice is linked to nuclear factor-κB. *J. Biol. Chem.* 277: 17821-17829.

Simon M., Seraphin B., and Faye G. 1986. KIN28, a yeast split gene coding for a putative protein kinase homologous to CDC28. *EMBO J.* 5: 2697-2701.

Sims R. J., 3rd, Belotserkovskaya R., and Reinberg D. 2004a. Elongation by RNA polymerase II: The short and long of it. *Genes Dev.* 18: 2437-2468.

Sims R. J., 3rd, Mandal S. S., and Reinberg D. 2004b. Recent highlights of RNA-polymerase-II-mediated transcription. *Curr. Opin. Cell Biol.* 16: 263-271.

Singh H., Erkine AM., Kremer S. B., Duttweiler H. M., Davis D. A., Iqbal J., Gross R. R., and Gross D. S. 2006. A functional module of yeast mediator that governs the dynamic range of heat-shock gene expression. *Genetics* 172: 2169-2184.

Smale ST. and Kadonaga J. T. 2003. The RNA polymerase II core promoter. *Annu. Rev. Biochem.* 72: 449-479.

Smith C. L. and Peterson C. L. 2005. ATP-dependent chromatin remodeling. *Curr. Top. Dev. Biol.* 65: 115-148.

Solow S., Salunek M., Ryan R., and Lieberman P. M. 2001. TAF_{II}250 phosphorylates human transcription factor IIA on serine residues important for TBP binding and transcription activity. *J. Biol. Chem.* 276: 15886-15892.

Sopta M., Burton Z. F., and Greenblatt J. 1989. Structure and associated DNA-helicase activity of a general transcription initiation factor that binds to RNA polymerase II. *Nature* 341: 410-414.

Spangler L., Wang X., Conaway J. W., Conaway R. C., and Dvir A. 2001. TFIIH action in transcription initia-

tion and promoter escape requires distinct regions of downstream promoter DNA. *Proc. Natl. Acad.* Sci. 98: 5544-5549.

Steger D. J., Owen-Hughes T., John S., and Workman I. L. 1997. Analysis of transcription factor-mediated remodeling of nucleosomal arrays in a purified system. *Methods* 12: 276-285.

Stevens J. L., Cantin G. T., Wang G., Shevchenko A., and Berk A. J. 2002. Transcription control by EIA and MAP kinase pathway via Sur2 mediator subunit. *Science* 296: 755-758.

Stuempfle K. J., Koptides M., Karinch AM., and Floros J. 1996. Preparation of transcriptionally active nuclear extracts from mammalian tissues. *BioTechniques* 21: 48-50, 52.

Sumimoto H., Ohkuma Y., Sinn E., Kato H., Shimasaki S., Horikoshi M., and Roeder R. G. 1991. Conserved sequence motifs in the small subunit of human general transcription factor TFIIE. *Nature* 354: 401-404.

Sun X., Ma U., Sheldon M., Yeung K., and Reinberg D. 1994. Reconstitution of human TFIIA activity from recombinant polypeptides: A role in TFIID-mediated transcription. *Genes Dev.* 8: 2336-2348.

Sun X., Zhang Y., Cho H., Rickert P., Lees E., Lane W., and Reinberg D. 1998. NAT, a human complex containing Srb polypeptides that functions as a negative regulator of activated transcription. *Mol. Cell* 2: 213-222.

Svejstrup J. Q. 2004. The RNA polymerase II transcription cycle: Cycling through chromatin. *Biochim. Biophys. Acta* 1677: 64-73.

Taatjes D. J., Nsar A. M., Andel F., 3rd, Nogales E., and Tjian R. 2002. Structure, function, and activator-induced conformations of the CRSP coactivator. *Science* 295: 1058-1062.

Tabtiang R. K. and Herskowitz I. 1998. Nuclear proteins Nutip and Nut2p cooperate to negatively regulate a Swi4p-dependent lacZ reporter gene in Saccharomyces cerevisiae. *Mol. Cell. Biol.* 18: 4707-4718.

Takagi Y., Masuda C. A., Chang W. H., Komori H., Wang D., Hunter T., Joazeiro C. A., and Kornberg R. D. 2005. Ubiquitin ligase activity of TFIIH and the transcriptional response to DNA damage. *Mol. Cell* 18: 237-243.

Tantin D. and Carey M. 1994. A heteroduplex template circumvents the energetic requirement for ATP during activated transcription by RNA polymerase II. *J. Biol. Chem.* 269: 17397-17400.

Tantin D., Chi T., Hon R., Pyo S., and Carey M. 1996. Biochemical mechanism of transcriptional activation by GAL4-VP16. *Methods Enzymol.* 274: 133-149.

Theron T., Fousteri M. I., Volker M., Harries L. W., Botta E., Stefanini M., Fujimoto M., Andressoo J. O., Mitchell J., Jaspers N. G., et al. 2005. Transcription-associated breaks in xeroderma pigmentosum group D cells from patients with combined features of xeroderma pigmentosum and Cockayne syndrome. *Mol. Cell. Biol.* 25: 8368-8378.

Thompson N. E. and Burgess R. R. 1996: Immunoaffinity purification of RNA polymerase II and transcription factors using polyol-responsive monoclonal antibodies. *Methods Enzymol.* 274: 513-526.

Thompson N. E., Aronson D. B., and Burgess R. R. 1990. Purification of eukaryotic RNA polymerase II by immunoaffinity chromatography. Elution of active enzyme with protein stabilizing agents from a polyol-responsive monoclonal antibody. *J. Biol. Chem.* 265: 7069-7077.

Timmers H. T. 1994. Transcription initiation by RNA polymerase II does not require hydrolysis of the β-γ phosphoanhydride bond of ATP. *EMBO J.* 13: 391-399.

Tirode F., Busso D., Coin F., and Egly J. M. 1999. Reconstitution of the transcription factor TFIIH: Assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Mol. Cell* 3: 87-95.

Toth J. 1., Datta S., Athanikar J. N., Freedman L. P., and Osborne T. F. 2004. Selective coactivator interactions in gene activation by SREBP-la and-1c. *Mol. Cell. Biol.* 24: 8288-8300.

Treisman J. 2001. *Drosophila* homologues of the transcriptional coactivation complex subunits TRAP240 and TRAP230 are required for identical processes in eye-antennal disc development. *Development* 128: 603-615.

Tschochner H., Sayre M. H., Flanagan P. M., Feaver W. J., and Kornberg R. D. 1992. Yeast RNA polymerase 11 initiation factor e: Isolation and identification as the functional counterpart of human transcription factor IIB. *Proc. Natl. Acad. Sci.* 89: 11292-11296.

Tsuboi A., Conger K., Garrett K. P., Conaway R. C., Conaway J. W., and Arai N. 1992. RNA polymerase II initiation factor a from rat liver is almost identical to human TFIIB. *Nucleic Acids Res.* 20: 3250.

Tubon T. C., Tansey W. P., and Herr W. 2004. A nonconserved surface of the TFIIB zinc ribbon domain plays a direct role in RNA polymerase II recruitment. *Mol. Cell. Biol.* 24: 2863-2874.

Tudor M., Murray P. J., Onufryk C., Jaenisch R., and Young R. A. 1999. Ubiquitous expression and embryonic requirement for RNA polymerase II coactivator subunit Srb7 in mice. *Genes Dev.* 13: 2365-2368.

Tutter A. V., Fryer C. J., and Jones K. A. 2001. Chtomatin-specific regulation of LEF-1-β-catenin transcription activation and inhibition in vitro. *Genes Dev.* 15: 3342-3354.

Uesugi M., Nyanguile O., Lu H., Levine A. J., and Verdine G. L. 1997. Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science* 277: 1310-1313.

Upadhyaya A. B., Khan M., Mou T. C., Junker M., Gray D. M., and DeJong J. 2002. The germ cell-specific transcription factor ALF. Structural properties and stabilization of the TATA-binding protein (TBP) -DNA complex. *J. Biol. Chem.* 277: 34208-34216.

van de Peppel J., Kettelarij N., van Bakel H., Kockelkorn T. T., van Leenen D., and Hoistege F. C. 2005. Mediator expression profiling epistasis reveals a signal transduction pathway with antagonistic submodules and highly specific downstream targets. *Mol. Cell* 19: 511-522.

Van Dyke M. W., Roeder R. G., and Sawadogo M. 1988. Physical analysis of transcription preinitiation complex assembly on a class II gene promoter. *Science* 241: 1335-1338.

Verdone L., Caserta M., and Di Mauro E. 2005. Role of histone acetylation in the control of gene expression. *Biochem. Cell Biol.* 83: 344-353.

Verrijzer C. P., Yokomori K., Chen J. L., and Tjian R. 1994. *Drosophila* $TAF_{II}150$: Similarity to yeast gene TSM-1 and specific binding to core promoter DNA. *Science* 264: 933-941.

Wang B. Q., Lei L., and Burton Z. F. 1994. Importance of codon preference for production of human RAP74 and reconstitution of the RAP30/74 complex. *Protein Expr. Purif.* 5: 476-485.

Wang B. Q., Kostrub CE, Finkelstein A., and Burton Z. E 1993. Production of human RAP30 and RAP74 in bacterial cells. *Protein Expr. Purif.* 4: 207-214.

Wang G., Balamotis M. A., Stevens J. L., Yamaguchi Y., Handa H., and Berk A. J. 2005. Mediator requirement for both recruitment and postrecruitment steps in transcription initiation. *Mal. Cell* 17: 683-694.

Wang W., Carey M., and Gralla J. D. 1992a. Polymerase II promoter activation: Closed complex formation and ATP-driven start site opening. *Science* 255: 450-453.

Wang W., Gralla J. D., and Carey M. 1992b. The acidic activator GAL4-AH can stimulate polymerase II transcription by promoting assembly of a closed complex requiring TFIID and TFIIA. *Genes Dev.* 6: 1716-1727.

Wang W., Xue Y., Zhou S., Kuo A., Cairns B. R., and Crabtree G. R. 1996. Diversity and specialization of mammalian SWI/SNF complexes. *Genes Dev.* 10: 2117-2130.

Wang Y., Li Y., Zeng W., Zhu C., Xiao J., Yuan W., Cai Z., Zhou J., Liu M., and Wu X. 2004. IXL, a new subunit of the mammalian Mediator complex, functions as a transcriptional suppressor. *Biochem. Biophys. Res. Commun.* 325: 1330-1338.

Warfield L., Ranish J. A., and Hahn S. 2004. Positive and negative functions of the SAGA complex mediated through interaction of Spt8 with TBP and the N-terminal domain of TFIIA. *Genes Dev.* 18: 1022-1034.

Wassarman D. A. and Sauer F. 2001. $TAF_{I}250$: A transcription toolbox. *J. Cell Sci.* 114: 2895-2902.

Watanabe T., Hayashi K., Tanaka A., Furumoto T., Hanaoka F., and Ohkuma Y. 2003. The carboxy terminus of the small subunit of TFIIE regulates the transition from transcription initiation to elongation by RNA polymer-

ase II. *Mol. Cell. Biol.* 23：2914-2926.

Weber C. A.，Salazar E. P.，Stewart S. A.，and Thompson L. H. 1990. ERCC2：cDNA cloning and molecular characterization of a human nucleotide excision repair gene with high homology to yeast RAD3. *EMBO J.* 9：1437-1447.

Weeda G.，van Ham R. C.，Masurel R.，Westerveld A.，Odijk H.，de Wit J.，Bootsma D.，van der Eb A. J.，and Hoeijmakers J. H. 1990. Molecular cloning and biological characterization of the human excision repair gene ERCC-3. *Mat. Cell. Biol.* 10：2570-2581.

Wei W.，Dorjsuren D.，Lin Y.，Qin W.，Nomura T.，Hayashi N.，and Murakami S. 2001. Direct interaction between the subunit RAP30 of transcription factor IIF（TFIIF）and RNA polymerase subunit 5，which contributes to the association between TFIIF and RNA polymerase II. *J. Biol. Chem.* 276：12266-12273.

Weideman C. A.，Netter R. C.，Benjamin L. R.，McAllister J. J.，Schmiedekamp L. A.，Coleman R. A.，and Pugh B. F. 1997. Dynamic interplay of TFIIA，TBP and TATA DNA. *J. Mol. Biol.* 271：61-75.

Well P. A.，Luse D. S.，Segall J.，and Roeder R. G. 1979a. Selective and accurate initiation of transcription at the Ad2 major late promotor in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* 18：469-484.

Weil P. A.，Segall J.，Harris B.，Ng S. Y.，and Roeder R. G. 1979b. Faithful transcription of eukaryotic genes by RNA polymerase III in systems reconstituted with purified DNA templates. *J. Biol. Chem.* 254：6163-6173.

Winkler G. S.，Vermeulen W.，Coin F.，Egly J. M.，Hoeijmakers J. H.，and Weeda G. 1998. Affinity purification of human DNA repairl transcription factor TFIIH using epitope-tagged xeroderma pigmentosum B protein. *J. Biol. Chem.* 273：1092-1098.

Woychik N. A. and Hampsey M. 2002. The RNA polymerase II machinery：Structure illuminates function. *Cell* 108：453-463.

Yamit-Hezi A. and Dikstein R. 1998. TAF_{II}105 mediates activation of anti-apoptotic genes by NF-κB. *EMBO J.* 17：5161-5169.

Yan Q.，Moreland R. J.，Conaway J. W.，and Conaway R. C. 1999. Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. *J. Biol. Chem.* 274：35668-35675.

Yang F.，DeBeaumont R.，Zhou S.，and Naar A. M. 2004. The activator-recruited cofactor/Mediator coactivator subunit ARC92 is a functionally important target of the VP 16 transcriptional activator. *Proc. Natl. Acad. Sci.* 101：2339-2344.

Yang X. J.，Ogryzko V. V.，Nishikawa J.，Howard B. H.，and Nakatani Y. 1996. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* 382：319-324.

Yang Z.，Yik J. H.，Chen R.，He N.，Jang M. K.，Ozato K.，and Zhou Q. 2005. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mat. Cell* 19：535-545.

Yankulov K. and Bentley D. 1998. Transcriptional control：Tat cofactors and transcriptional elongation. *Curr. Biol.* 8：R447-449.

Yee A.，Nichols M. A.，Wu L.，Hall F. L.，Kobayashi R.，and Xiong Y. 1995. Molecular cloning of CDK7-associated human MAT1，a cyclin-dependent kinase-activating kinase（CAK）assembly factor. *Cancer Res.* 55：6058-6062.

Yoda A.，Kouike H.，Okano H.，and Sawa H. 2005. Components of the transcriptional Mediator complex are required for asymmetric cell division in *C. elegans*. *Development* 132：1885-1893.

Yokomori K.，Admon A.，Goodrich J. A.，Chen J. L.，and Tjian R. 1993. *Drosophila* TFIIA-L is processed into two subunits that are associated with the TBP/TAF complex. *Genes Dev.* 7：2235-2245.

Yokomori K.，Zeidler M. P.，Chen J. L.，Verrijzer C. P.，Mlodzik M.，and Tjian R. 1994. *Drosophila* TFIIA directs cooperative DNA binding with TBP and mediates transcriptional activation. *Genes Dev.* 8：2313-2323.

Yoon H.，Miller S. P.，Pabich E. K.，and Donahue T. F. 1992. SSL1，a suppressor of a HIS4 5′-UTR stem-loop

mutation, is essential for translation initiation and affects UV resistance in yeast. *Genes Dev.* 6: 2463-2477.

Yu X., Li P., Roeder R. G., and Wang Z. 2001. Inhibition of androgen receptor-mediated transcription by amino-terminal enhancer of split. *Mol. Cell. Biol.* 21: 4614-4625.

Yudkovsky N., Ranish J. A., and Hahn S. 2000. A transcription reinitiation intermediate that is stabilized by activator. *Nature* 408: 225-229.

Zakharova N., Lymar E. S., Yang E., Malik S., Zhang J. J., Roeder R. G., and Darnell J. E., Jr. 2003. Distinct transcriptional activation functions of STAT1α and STAT1β on DNA and chromatin templates. *J. Biol. Chem.* 278: 43067-43073.

Zawel L., Kumar K. P., and Reinberg D. 1995. Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev.* 9: 1479-1490.

Zenzie-Gregory B., O' Shea-Greenfield A., and Smale S. T. 1992. Similar mechanisms for transcription initiation mediated through a TATA box or an initiator element. *J. Biol. Chem.* 267: 2823-2830.

Zerby D. and Lieberman P. M. 1997. Functional analysis of TFIID activator interaction by magnesium-agarose gel electrophoresis. *Methods* 12: 217-223.

Zhang C., Zobeck K. L., and Burton Z. F. 2005. Human RNA polymerase II elongation in slow motion: Role of the TFIIF RAP74 α1 helix in nucleoside triphosphate-driven translocation. *Mol. Cell. Biol.* 25: 3583-3595.

Zhang F., Sumibcay L., Hinnebusch A. G., and Swanson M. J. 2004. A triad of subunits from the Galil/tail domain of Srb mediator is an in vivo target of transcriptional activator Gcn4p. *Mol. Cell. Biol.* 24: 6871-6886.

Zhou Q., Lieberman P. M., Boyer T. G., and Berk A. J. 1992. HoloTFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Genes Dev.* 6: 1964-1974.

Zhu X., Wiren M., Sinha I., Rasmussen N. N., Linder T., Holmberg S., Ekwall K., and Gustafsson C. M. 2006. Genome-wide occupancy profile of mediator and the Srb8-11 module reveals interactions with coding regions. *Mol. Cell* 22: 169-178.

Zurita M. and Merino C. 2003. The transcriptional complexity of the TFIIH complex. *Trends Genet.* 19: 578-584.

网络资源

http://www.fhcrc.org/science/labs/hahn：Hahn 实验室网站提供与 RNA Pol II 和 Pol III 真核转录机制研究相关的方法、参考文献及其他资源。

http://www.nccc.com：国家细胞培养中心（The National Cell Culture Center）为研究团体提供定制性的大规模细胞培养服务。

13　体外研究染色质动力学：染色质组装、重塑和转录

要点

- 内源性染色质结构可以在体外重现吗？
- 当尝试在体外重现特定染色质结构时，关键的选择是模板、组蛋白来源和核小体组装方法。
- 可采取几种方法来评价染色质重建的完整性和质量。
- 固定化模板分析可剖析各种染色质“重塑/修饰”酶在体外转录调控中的作用。
- ATP 依赖性染色质重塑酶具有可影响转录活性的不同生化活性。
- 位点特异性化学可以改变重组组蛋白，生成分析染色质结构和功能的新试剂。

引言和概述

正如前面章节中所述，在真核基因调控领域一个突出的问题是理解基因特异性激活因子调控转录的生化机制。在第 11 章和第 12 章中，我们讨论了剖析这些转录调控因子如何与其同源 DNA 序列相互作用和从 DNA 模板上增强转录的生化策略。但考虑到细胞内真核基因组被组装成核小体阵列，完整重现控制复杂基因调控事件的通路，还需要在重建成染色质的 DNA 模板上进行机制研究。

正如第 1 章中所述，核小体组装和染色质纤维的形成，为导致体内外转录起始和延伸的通路的几乎每一步都制造了强有力的障碍。结果，染色质动力学的调控在控制基因

表达中起着不可或缺的作用。在大多数情况下，基因特异性激活因子或抑制因子在这一进程中发挥核心作用，它们协调大量染色质重塑和修饰酶的募集，而这些酶在染色质环境中调控转录。三磷酸腺苷（ATP）依赖性重塑酶能改变核小体位置、驱逐组蛋白八聚体、取代 H2A/H2B 二聚体或通过引入或消除组蛋白变体改变染色质组成。在一个特定的靶基因上可能使用这些活性的全部或特定子集。同样，还有大量基因特异性调控因子指导的位点特异性组蛋白修饰，但这些组蛋白修饰如何在基因调控通路的不同步骤发挥作用基本上还不清楚。因此，发展重现内源性染色质结构和转录相关核小体重排及修饰完整图谱的体外系统，对于生化机制的详细理解具有决定性的作用。

本章的目的是描述用于组装和分析染色质重建的策略，以便用于随后的分析，在转录分析方法、因子结合和（或）染色质重塑/修饰酶分析方法中都需要这些分析。本章还描述了染色质重塑和组蛋白修饰酶的分析方法，以作为染色质模板体外转录分析的策略；讨论了每种分析的优点和缺点，以及研究人员对研究中可能出现的技术问题的警觉，每个策略在时间和资源投入上亦有不同。我们的总体目标是为研究人员提供剖析一个生理性模板，即染色质上复杂的体外转录通路的每个步骤的所有实验工具和策略，在规划一项需要染色质重建的研究时，一般需考虑如下几个问题。

• 用什么作核心组蛋白的来源？组蛋白是丰富的蛋白质，因此，它们可以从多种来源的细胞中纯化。然而，重组组蛋白的优点是它们缺乏所有的翻译后修饰，并且它们的序列可以通过改变克隆的基因而得到操控。但是，纯化重组组蛋白需要大量的精力和资源。尽管组蛋白在真核生物中高度保守，但序列变异的确存在（特别是组蛋白 H2A 和 H2B），因此，组蛋白的细胞来源（如酵母或人类）在开始一项研究之前应予以考虑。

• 该研究需要单核小体或核小体阵列的组装吗？单核小体长度的 DNA（150～220 bp）可能足以包含核心启动子及启动子近端调控元件。这些简单的底物可能适合分析因子与核小体位点的结合或组蛋白修饰或重塑酶活性的研究，或者适合于研究转录起始或延伸的最初步骤。显然，单核小体在染色质武器库中是重要的工具，但是必须始终牢记，单核小体在细胞中并不存在。的确，在缺乏相邻核小体的情况下，组蛋白氨基末端结构域显著重排（Usachenko et al. 1994），并且 H4 的氨基末端结构域封阻了核小体二分体轴上的 DNA 的可及性（Vettese-Dadey et al. 1994，1996）。重建核小体阵列更加接近地模拟生理染色质，因此，如果一个项目的长远目标是在体外重建特定的调控事件，那么这些底物是最有用的。

• 染色质组装是使用纯化的组分呢，还是一个粗制组装系统就足够了？细胞质粗提物可将大的 DNA 分子有效地组装成均匀间隔的核小体阵列，但是，这些提取物含有丰富的染色质重塑酶，会使进一步的分析复杂化。在这种情况下，可能需要染色质模板的广泛纯化。与此相反，盐透析方法相当简单，只需要组蛋白和 DNA。虽然这种方法很适合单核小体组装，但其重建产生非常紧密排列的、间距不规则的核小体。因此，使用这种方法，较长 DNA 模板上核小体的生理定位需包含核小体定位序列。

实验策略

用于组装染色质的策略

将纯化的组蛋白和 DNA 组装成染色质样颗粒的方法首先发展于 20 世纪 70 年代初

(Oudet et al. 1975)，时间接近 Thomas 和 Kornberg 正式发现核小体 (Kornberg 1974; Kornberg and Thomas 1974)。早期的方法是使用透析法将来自高盐或高尿素的组蛋白-DNA 混合物透析成低盐缓冲液。这种盐透析重建方法在现阶段仍然是染色质重建最简单、应用最广泛的方法之一。随后，Laskey 等 (1977) 从非洲爪蟾未受精卵中发展了 ATP 依赖性无细胞系统，该系统可用周期性、生理间隔的核小体 (约 200 bp) 组装核小体阵列。随后报道了使用爪蟾卵母细胞质提取物 (Glikin et al. 1984) 和果蝇胚胎提取物的类似的系统 (Becker et al. 1994; Kamakaka and Kadonaga 1994)。最近，Kadonaga 小组纯化了果蝇胚胎提取物系统的活性成分，他们发展了完全重组的组装系统，该系统在一个 ATP 依赖性反应中产生长的、周期型核小体阵列 (Ito et al. 1999; Fyodorov Kadonaga 2003)。在下面的一节中，我们将介绍这些不同的方法并讨论它们的优缺点。

通过盐透析的染色质组装

组蛋白八聚体在含有高盐 (即 2 mol/L NaCl) 的缓冲液中是一种稳定的低聚复合物，从而反映了一个事实，即推动八聚体组装的分子相互作用中，疏水相互作用占主导地位 (Arents et al. 1991; Luger et al. 1997)。组蛋白八聚体的高盐稳定性促成了 Moudrianakis 研究组对其分辨率为 3.1Å 的 X 射线结构的确定 (Arents et al. 1991)。将八聚体转移进低盐 (0～200 mmol/L NaCl) 缓冲液中会导致八聚解离成两个 H2A/H2B 二聚体和一个 H3/H4 四聚体。这些低聚物要进一步解离成单个的组蛋白需要极端 pH 或高浓度尿素的处理 (Eickbush and Moudrianakis 1978)。组蛋白是高电荷的蛋白质，因此在低盐缓冲液中它们倾向于黏附玻璃和塑料管，因此，当使用组蛋白时，建议用硅化管。

当 DNA 存在时，如果一个组蛋白八聚体溶液被迅速稀释成低盐，就会形成非特异性聚集。不过，如果将一个 DNA 和组蛋白八聚体溶液通过梯度透析、逐步透析或逐步稀释 (也称为盐跳)，缓慢稀释到低盐条件，则可组装出真正的核小体。借助盐透析的染色质组装利用的是盐浓度低于 2 mol/L NaCl 时组蛋白八聚体的拆卸，以及 H3/H4 四聚体和二聚体 H2A/H2B 独特的 DNA 结合特性。例如，当盐浓度从 2 mol/L 下降到 1 mol/L 时，八聚体拆卸，并且 H3/H4 四聚体结合到 DNA 上，组织约 90 bp。在这种盐浓度下，H2A/H2B 二聚体不能稳定地结合 DNA。当盐浓度进一步降低时，H2A/H2B 二聚体结合到四体颗粒，当透析到 0.6 mol/L NaCl 时，则形成完整的组装 (Hansen et al. 1991)。之后，组装物透析或稀释成低盐缓冲液 (0.05～0.2 mol/L NaCl)，在这种状态下，染色质在冰上可稳定保存数周 (不要冻结)。

单核小体组装往往包含一些游离 DNA 或聚集的组蛋白-DNA 复合物，并且在此情况下，单核小体可以很容易地通过在线性蔗糖梯度中的沉降加以纯化。核小体阵列重建一般不需要纯化。盐透析重建方法的支持者认为，用这种方法逐步组装核小体可能可以很接近地模拟体内复制叉后核小体的组装，因为该处在加入 H2A/H2B 二聚体之前，H3/H4 四聚体被组装到子链 (daughter strand) 上。

除了使用盐透析将纯化组蛋白沉积到 DNA 上外，也可以使用高盐浓度将来自细胞染色质片段的核小体“捐赠”到 ^{32}P 标记的 DNA 片段上。该方法的试剂很容易准备，

它不需要纯化的组蛋白，并且一般情况下，我们从来不用纯化用这种方法组装的单核小体。Workman 小组也用此捐赠者转移的方法组装小批量放射性核小体阵列，以便用于转录分析和组蛋白乙酰转移酶复合物分析（Steger et al. 1998）。这种方法需要通过有限的微球菌核酸酶（MNase）消化和凝胶过滤分离，从孤立的细胞核中准备短的寡核小体（4～10 聚体）（Owen-Hughes et al. 1999）。纳克级的^{32}P 标记 DNA 片段（通常单核小体长度）与大摩尔过量的寡核小体混合，并将盐浓度提高到 2 mol/L NaCl，以解离组蛋白八聚体。然后该反应相继用含有较低盐浓度的缓冲液稀释，并且八聚体在标记探针和细胞 DNA 之间重新分配。通过改变探针 DNA 与寡核小体的比例，可将探针100%地组装成单核小体。

这种八聚体转移技术的一个缺点是随后的分析将包含细胞寡核小体。需要注意的是，用蔗糖梯度离心纯化标记的单核小体会产生非常稀的染色质溶液，这种稀溶液不是很稳定。一个已经被 Workman 研究组成功使用的替代策略是，在一个生物素化的 DNA 片段上组装单核小体（或核小体阵列）。由此产生的生物素化单核小体可以通过链亲和素修饰琼脂糖微珠捕获的方式从细胞寡核小体中纯化出来。单核小体可由限制性内切核酸酶消化从珠子上释放在小体积溶液中。

优点和缺点。核小体重建的盐透析方法与其他重建方法相比有几个优势。首先，这种方法只需要很少几种试剂，即纯化的组蛋白和 DNA。此外，可以很容易地通过改变组蛋白八聚体与 DNA 的比例来控制长 DNA 片段上核小体的密度。举例来说，大约每 160 bp DNA 一个八聚体的比例将产生 DNA 被核小体完全加载（饱和）的组装，而使用更少的组蛋白八聚体将产生亚饱和染色质片段。然而，重要的是要认识到尽管 DNA 片段能趋于被核小体饱和（即每 160 bp 一个核小体），但组蛋白八聚体的沉积并未显示出它是一种饱和反应（关于讨论，见 Carruthers et al. 1999）。因此，如果采用高比率的组蛋白八聚体与 DNA，大量的八聚体将沉积到 DNA 上，但这并不代表真正的核小体。这些非特异性结合的八聚体会干扰随后的反应，同时它们也能导致低浓度二价阳离子（如 Mg^{2+}）存在下核小体阵列的聚集。由于这个原因，在每个重建反应的组装之前，应确定组蛋白八聚体和 DNA 的浓度，以便计算精确的比值。注意，一个含有甚至 10%过量组蛋白八聚体的组装有时会导致不成比例的高水平聚集。

盐透析重建在几个应用中未必是染色质组装的最好方法。当八聚体经盐透析沉积在 DNA 长片段上时，由此产生的核小体往往以非生理间距紧密包装（见专题 13.1）。可变长度的无核小体 DNA 也可能存在在紧密包装的核小体补丁之间。因此，如果目标是重建定位在一个 DNA 长片段上的生理性相关核小体，盐透析重建可能并不适宜。避免这个问题的一个策略是使用模式核小体阵列模板，在这种模板上，一个目标 DNA 片段（200～400 bp）被插入串联的、头尾相接的核小体定位元件（NPE）(如 5S rRNA 基因）重复之间（Simpson et al. 1985）（专题 13. 1）。在核小体组装期间，5S 基因上核小体的定位强制定位的核小体在插入的调控区上的组装。Workman 及其同事已经用这种方法来组装核小体阵列，在这种阵列上，核小体被定位在人类免疫缺陷病毒 1 型（HIV-1）的启动子上（Steger and Workman 1997）或包含了 Gal4p-VP16 激活因子 5 个结合位点的腺病毒 E4 启动子上（Utley et al. 1998）。这些模板被成功用于监测激活因子依赖的

组蛋白乙酰转移酶复合物募集（Utley et al. 1998；见图 13.9，下文）、打靶 SWI/SNF 染色质重塑复合物（Yudkovsky et al. 1999），以及体外转录（Steger et al. 1998；Utley et al. 1998；Ikeda et al. 1999；Neely et al. 1999）。

专题 13.1 DNA 模板和重建方法对阵列内核小体定位的影响

细胞内散装染色质（bulk chromatin）一般由长的核小体阵列构成，阵列内的核小体每隔一定间隔分开（图 13.1）。第一种策略是，当组蛋白八聚体通过盐透析方法沉积到 DNA 模板上时，会产生一种非生理型的、密集包装的核小体排列，这种结构可在阵列内留下大的游离 DNA 区域（图 13.1，左下图）。第二种策略是，使用带有头尾相接重复 NPE（如海胆 5S rDNA 基因）的 DNA 模板。如果一组 NPE 位于目标序列（图中显示的一个含有腺病毒 E4 启动子的二核小体长度的片段）的侧翼，由 NPE 元件定位的核小体会创建指导核小体定位于中心启动子片段上的边界（图 13.1，上图）。通过改变中心片段上启动子的位置，可控制启动子元件的核小体定位。第三种策略是，ATP 依赖性 ACF/NAP1 组装法用于在异源 DNA 片段上重建均匀间隔的核小体阵列（图 13.1，右下图）。请注意，在这种情况下，核小体均匀间隔，但群体内每个阵列有不同的一组转录位置。核小体的随机定位导致在部分阵列内出现可及的启动子区域（注意箭头位置）。

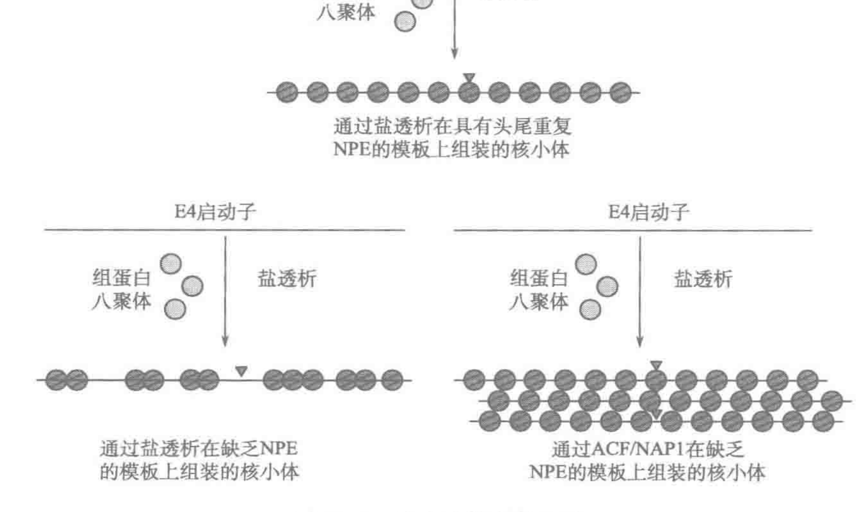


图 13.1 核小体阵列的重建。

体内研究可能表明，一个或更多的序列特异性基因调控蛋白在转录激活前结合到无核小体区域。为了在体外组装这种类型的结构，需要在核小体组装前预结合这些因子。

然而，暴露于 0.35 mol/L 以上的盐浓度往往会从其位点上剥离这些因子。对于这些类型的分析，必须使用核小体组装的替代方法，在这些替代方法中，核小体在低盐缓冲液中可被沉积。一些 DNA-蛋白质复合物是显著抗盐的，并因此与盐透析重建兼容。最值得注意的是，Luse 及其同事（Chang and Luse 1997）在通过快速盐透析组装质粒染色质之前，在质粒 DNA 上预制 RNA 聚合酶 II（RNAPII，或 Pol II）延伸复合物。这些研究最早充分论证了核小体对转录延伸的抑制作用。

最后，根据我们的经验，当使用酵母组蛋白八聚体组装核小体阵列时，盐透析重建并不是首选方法。使用这种方法时，即使有了重组酵母组蛋白或从酵母细胞核纯化的酵母组蛋白八聚体，核小体阵列重建的效率和可重复性也非常低。Laybourn 组也观察到了类似的现象（Wongwisansri and Laybourn 2004）。这些效应的分子基础尚不清楚，但这种分子基础有可能是在较高盐的浓度下，酵母 H2A/H2B 二聚体比脊椎动物组蛋白更好地结合 DNA，并因此干扰正确的逐步组装。与此相反，Wu 实验室已报道了快速盐稀释方法，并成功地用于酵母单核小体的组装（Shen et al. 2003）。

使用组蛋白伴侣的染色质组装

在生理盐浓度下，将组蛋白和 DNA 直接混合会导致 DNA-组蛋白聚集体的形成。通过在反应中包含带负电荷的“组装因子”，可以避免这种反应，这些因子都被称为组蛋白伴侣，但本质上它们只是组蛋白结合因子。从历史上看，多聚谷氨酸是第一个用于核小体组装的带负电的聚合物（Stein and Mitchell 1988）。随后，酸性组蛋白结合蛋白，如 NAP1（Ishimi and Kikuchi 1991；Ito et al. 1996），被确定还能够在生理盐缓冲液中组装核小体。

一般情况下，将纯化的组蛋白八聚体（在高盐中）与过量的伴侣混合（对于 yNAP1，比例是 4∶1），然后在低盐缓冲液中透析。之后，组蛋白-伴侣复合物与目标 DNA 孵育几个小时，以完成核小体组装。当将组蛋白-伴侣复合物加入 DNA 中时，八聚体由伴侣转移到 DNA 上，这似乎是在一个单步反应中完成的，从而产生一个组装的核小体和游离伴侣。对于 yNAP1，用线性或环状 DNA 模板组装都是有效的，而果蝇和小鼠 NAP1 在用超螺旋质粒 DNA 时才能发挥最佳功能。由于 yNAP1 是一个具有约 90 kDa 自然分子质量的二聚体，通常通过蔗糖梯度沉降从核小体组装物中纯化这种蛋白质。

优点和缺点。伴侣介导的核小体组装的主要功用是其在生理盐条件下组装核小体的能力。Laybourn 及其同事已经使用 yNAP1 在包含 PHO5 启动子的质粒 DNA 上组装了酵母八聚体，并且在这种情况下，他们能够在启动子元件上重建生理性的核小体定位（Terrell et al. 2002）。的确，对于利用酵母组蛋白的重建，伴侣依赖性组装似乎是首选方法。但是，与盐透析重建情况类似，伴侣介导的组装通常在缺乏核小体定位序列的长 DNA 片段上产生紧密排列、间隔不均匀的核小体。此外，可能很难从被组装的核小体中完全去除组蛋白伴侣。污染的伴侣可能会干扰后续分析。对于 yNap1 来说这可能尤其成问题，因为它足以催化 H2A/H2B 二聚体从核小体上丧失，并增强核小体的流动性（Park et al. 2005）。

ATP 依赖性染色质组装

如上所述，盐透析和伴侣介导的核小体组装方法最大的缺点是，长 DNA 片段或质

粒上核小体的沉积不会导致周期性阵列，相反却导致密集排列的非均匀间隔的核小体。然而，有几个无细胞系统却能将质粒组装成染色质模板，这些模板上核小体以大约 200 bp 的生理性区间均匀间隔。这些组装系统是由组蛋白伴侣和 ATP 依赖性染色质重塑活性构成的，两者结合起来共同沉积和间隔核小体，即使在非常大的质粒中也是如此。早期的研究使用来自爪蟾卵母细胞（Glikin et al. 1984）或果蝇胚胎（Becker et al. 1994；Kamakaka and Kadonaga 1994）的粗制细胞质提取物（S150 或 S190）。在这些提取物中组装的质粒染色质已被广泛用于转录调控的生化研究。然而，大多数研究人员并不能获取足够数量的蛙卵母细胞或果蝇胚胎，而且从粗提物中纯化质粒染色质可能有问题。因此，目前不推荐这些类型的组装提取物。

1997 年，Kadonaga 及其同事报道了 ACF 的纯化，ACF 是果蝇胚胎提取物核小体组装系统的 ATP 利用成分（Ito et al. 1999；Fyodorov and Kadonaga 2003），是一种由 Acf1 和 ISWI ATP 酶组成的二亚基酶。这两个亚基的基因已被克隆到杆状病毒载体中，该载体可以将两个亚基单独或组合表达为 FLAG 标记多肽（见方案 13.3）。这些构体一般可从 Kadonaga 实验室获取。含有质粒 DNA、ACF、组蛋白伴侣（NAP1 或 Caf1）、核心组蛋白和 ATP 的反应混合物足以程序性指导长的周期性核小体阵列在生理盐条件下的组装（见方案 13.3）。要通过 NAP1 伴侣限制核小体的 ATP 非依赖性沉积，首先将质粒 DNA 通过真核拓扑异构酶 I（Topo I）处理，使其变得松弛。Topo I 可能也被包括在组装反应中，以消除这种核小体组装期间积累的正超螺旋。这些组装反应相当强大；ACF 可以以每 50 个组蛋白八聚体：1 个 ACF 的最佳比例程序性指导核小体组装。

优点和缺点。ATP 依赖性染色质组装产生酷似散装细胞染色质的周期性核小体阵列。当要将 DNA 大片段或质粒 DNA 重建成染色质时，这是首选方法。因为组装是在低盐缓冲液中进行，在组装反应期间可加入序列特异性 DNA 结合蛋白。这些蛋白质可以引起局部的核小体定位和模拟内源性座位的染色质结构的形成（Barton et al. 1993；Pazin et al. 1997）。许多研究者使用基于 ACF 的组装系统组装用于转录调控研究的质粒染色质（Jiang et al. 2000；An et al. 2004；An and Roeder 2004；Angelov et al. 2004；Guermah et al. 2006）。

ACF 依赖性组装方法的缺点之一是它需要多个组分的分离纯化及功能分析。特别是，ACF 的活性必须仔细滴定，因为高浓度实际上扰乱了核小体阵列的间距。一旦获得组装，研究者还必须决定是否从组装反应纯化染色质。污染性组装成分可能为后续分析的解释带来问题，因为 NAP1 可催化组蛋白 H2A/H2B 二聚体置换，并且 ACF 具有 ATP 依赖性染色质重塑酶功能。但是，大多数研究没有纯化染色质，研究者只是简单地承认有 ACF 及 NAP1 活性对观察到的结果发挥作用的可能性。一般来说，这种组装方法比盐透析组装需要投入更多的时间和资源。

虽然 ATP 依赖性组装方法生成周期性核小体阵列，但它没有组装出定位的核小体（见专题 13.1）。因此，在染色质分子的群体内，相对于特定 DNA 序列，核小体会随机定位。这与重建在含有头尾相连重复核小体定位序列的 DNA 模板上的定位阵列形成鲜明对比（专题 13.1）。随机定位的一个后果是，转录因子结合位点、限制性内切核酸酶识别位点、启动子元件等常会位于核小体之间的连接区域。因此，核小体组装可能不能

抑制转录过程中的几个步骤，而这些步骤可能在体内被定位的核小体封闭。此外，由于在很多体外转录分析中模板利用率会小于1%，所以在解释结果时需要考虑这些警告。

连接组蛋白的掺入

除了均匀间隔的核小体，来自后生动物物种的生理性染色质在每个核小体核心上还含有大约一分子的连接组蛋白。最近的研究表明，连接组蛋白与核小体核心颗粒的比值实际上随不同的组织而有变化，如在胚胎干细胞中比值低至0.5，而在胸腺细胞中为0.83（综述见 Fan et al. 2003，2005；Woodcock et al. 2006）。不太复杂的真核生物，如芽殖酵母（*Saccharomyces cerevisiae*），具有极低水平的连接组蛋白（每核心颗粒估计为0.03～0.25）。在一般情况下，连接组蛋白掺入的水平可能需要加以改变，以取得理想的生理性染色质结构。

无论是在盐透析或八聚体转移重建期间（Juan et al. 1994），还是在通过 ACF/NAP1 系统进行 ATP 依赖型组装期间（Fyodorov and Kadonaga 2003），连接组蛋白都可以被掺入。如果使用盐透析重建，则须在盐浓度降低到 0.6 mol/L NaCl 时，再将连接组蛋白加入重建反应。加入连接组蛋白后，透析持续进行到低盐浓度。在依靠 ACF/NAP1 进行 ATP 依赖型组装的情况下，将连接组蛋白简单地随核心组蛋白加入就可以了。重要的是，包含连接组蛋白的核小体阵列在含二价阳离子的缓冲液中相当容易自我缔合或“聚集”，因此，应通过快速微量离心仔细监测核小体阵列在不同缓冲溶液中的溶解度。

组蛋白来源

核心组蛋白

除粗提物系统外，所有推荐的染色质组装方法都需要外源性的纯化的组蛋白。组蛋白在真核生物中非常保守，鼠、鸡、人、苍蝇（果蝇）、青蛙（非洲爪蟾）中的组蛋白序列几乎相同（Marino-Ramirez et al. 2006）。因此，这些重组表达的组蛋白或纯化自这些细胞源的组蛋白中的每一个组蛋白都可视为是可互换的。与此相反，来自芽殖酵母[酿酒酵母（*S. cerevisiae*）]、裂殖酵母（*S. pombe*）或植物［拟南芥（*Arabidopsis*）］的组蛋白有明显的差异，特别是组蛋白 H2A 和 H2B。生化上，来自所有来源的组蛋白都能组装成组蛋白八聚体，这些组蛋白八聚体组装成约 147 bp 的 DNA，但这种组蛋白八聚体的稳定性却有所不同。众所周知，用重组酵母组蛋白组装的组蛋白八聚体和单核小体不及重组脊椎动物组蛋白组装的组蛋白八聚体和单核小体稳定，因此，组装原液应在组装后一周内使用。相比之下，植物八聚体似乎甚至比脊椎动物组蛋白更稳定，以致将植物八聚体解离成二聚体和四聚体时需要使用比通常用于其他来源八聚体的更低的盐浓度（小于或等于 0.6 mol/L NaCl）（Moehs et al. 1992），这可能给通过盐透析重建植物染色质带来麻烦。

历史上，组蛋白首次是通过细胞核的酸抽提分离的（Phillips and Johns 1965），此方法仍在使用，它可以产生单独的 H2A/H2B 二聚体和 H3/H4 四聚体（Chang and Luse 1997）。我们通常先从鸡红细胞或组织培养的 HeLa 细胞制备细胞核，再从细胞核中释放溶解的染色质，最后从溶解的染色质中纯化组蛋白（见方案 13.1）。简单地说，

细胞核用微球菌核酸酶处理，释放长的寡核小体，这些染色体片段与 CM-Sephadex 树脂在 0.35 mol/L NaCl 中混合，剥离寡核小体的连接组蛋白。剥离后的寡核小体用进一步的 MNase 消化修剪成较小的碎片，并且用羟基磷灰石（HTP）树脂层析从细胞 DNA 中纯化出组蛋白八聚体。也可以用 HTP 柱的梯度洗脱从 H3/H4 四聚体中分离 H2A/H2B 二聚体（Simon and Felsenfeld，1979）。注意，短的、由第二次 MNase 消化产生的寡核小体也可作为供体用于八聚体转移重建。一般情况下，我们从 200 ml 鸡血中可获得 50 mg 纯化的组蛋白八聚体。我们制备的鸡组蛋白八聚体溶液在 4℃可稳定保存一年多。

使用从散装染色质分离的组蛋白时需要注意的是，它们拥有一连串的翻译后修饰，可能使后续分析复杂化（Garcia et al. 2007）。此外，没有任何手段从纯化的组蛋白中消除修饰，这些修饰的丰度有时可通过以酶抑制剂处理细胞来调节。例如，丁酸钠是一种有效的组蛋白去乙酰化酶抑制剂，它经常被用于增加从 HeLa 细胞分离的组蛋白中的赖氨酸乙酰化水平（Tse et al. 1998）。

此外，单个组蛋白可在细菌中表达，重组组蛋白缺少翻译后修饰。通常，通过凝胶过滤和离子交换色谱，可将每个组蛋白从细菌包涵体中纯化为变性的多肽（Luger et al. 1999）。Tsukiyama 及其同事发表了该方案的修改版本，在修改后的方案中，使用连续的 MonoQ 和 MonoS 色谱在含尿素的缓冲液中纯化单个组蛋白（Vary et al. 2004）。之后，按等摩尔比混合所有 4 种变性核心组蛋白，并且通过盐浓度为 2 mol/L NaCl 的透析缓冲液重折叠组蛋白和重建八聚体。然后，组蛋白八聚体用凝胶过滤色谱法纯化。一种纯化和分析重组组蛋白的详细方案已经发表（Luger et al. 1999）。在我们的实验中，通过 Luger 法组装的重组八聚体不如自然组蛋白八聚体稳定，并且我们一般将它们在冰上只储存 1～2 个月。八聚体也可透析在含 50%甘油的缓冲液中，并在−80℃保存。

Gloss 及其同事已报道，在重折叠反应中，组蛋白多肽倾向于形成螺旋聚集（Gloss and Placek 2002）。为了减少这种反应，提高八聚体重建的整体效率，这些作者已发展出改进版的八聚体组装方法，即在八聚体组装前，将 H3/H4 四聚体和二聚体 H2A/H2B 作为独立的单元重折叠并组装。这种方法的关键是将变性的组蛋白快速稀释到缺乏变性剂的缓冲液中。在某些情况下，这种方法非常有效，以致不需要通过凝胶过滤层析纯化八聚体（L. Gloss，个人通讯）。虽然还没有测试，但这种八聚体重建的方法可能会产生更加稳定的重组酵母八聚体。

使用重组组蛋白比使用从细胞染色质获得的组蛋白有众多优势。当然，重组组蛋白翻译后修饰的缺乏使得它们成为体外转录分析的首选试剂。但首要的是，重组克隆可以通过诱变策略按意愿操控（见第 10 章）。可将组蛋白表达为含有氨基酸替换的组蛋白，这些氨基酸替换可能已经通过遗传研究被鉴定，还可以将氨基酸替换为之后要在体外进行分析的氨基酸。例如，酵母遗传学鉴定了单氨基酸替代的组蛋白 H3 和 H4 的等位基因（Sin 等位基因），这种单氨基酸替代的组蛋白缓解了转录对 SWI/SNF 重塑复合物的需求（Kruger et al. 1995）。这些组蛋白改变一直以来被假设可能创建一种模仿重塑状态的染色质结构。我们（Horn et al. 2002b）及其他人（Flaus et al. 2004）通过创造 Sin^{-} 重组组蛋白和组装单核小体及核小体阵列测试了这个想法。值得注意的是，发现

Sin⁻单核小体比野生型核小体在更低的温度下沿 DNA 滑动，而且 Sin⁻核小体阵列无法折叠成 30 nm 纤维。这些数据导致一种提议，即 SWI/SNF 可能在体内控制核小体的流动性和染色质折叠。

除了测试组蛋白残基的功能重要性外，可将新的半胱氨酸残基加到组蛋白上，以便合成后加入交联剂（Lee et al. 1999；Kan and Hayes 2007）、荧光化合物（Bruno et al. 2003b）或亲和纯化剂（如生物素）（Sinha and Peterson 2008）等。正如在下面的详细讨论，工程化的半胱氨酸残基也让研究者利用化学方法引入位点特异性的组蛋白翻译后修饰。最后，我们非常成功地使用变性的重组组蛋白作为抗原，生产未修改组蛋白的多克隆抗体。重组组蛋白唯一的缺点是，它们的大规模制备需要耗费大量的时间和资源。

连接组蛋白和组蛋白变体

已经可以从多种来源纯化连接组蛋白，包括鸡红细胞、培养的哺乳动物细胞（HeLa）和果蝇胚胎。在大多数情况下，可以在纯化核心组蛋白期间一步分离连接组蛋白（见方案 13.1）。例如，可以从纯化鸡八聚体期间使用的 CM-Sephadex 树脂中，利用盐梯度洗脱纯化连接组蛋白 H1 或 H5（一个红细胞特异的 H1 变体），或者可以在纯化 HeLa 核心组蛋白期间，从 HTP 柱的流穿馏分中分离部分纯化的连接组蛋白（Workman et al. 1991）。Kadonaga 及其同事也发表了一个使用疏水相互作用层析法纯化连接组蛋白 H1 的通用方法（Croston et al. 1991）。

在多数情况下，自然连接组蛋白制剂适合染色质重建和体外分析。但是，这些制剂几乎总是几个不同连接组蛋白变体或者亚型的混合物。小鼠有 6 种体细胞连接组蛋白 H1 的亚型，H1a～H1e 和 H1（0）（Woodcock et al. 2006）。每种 H1 亚型都有相似的整体结构，并且每种都结合核小体和维持高级染色质结构，但是，每种变体显示出有不同系列的翻译后修饰和组织分布（Woodcock et al. 2006；Wisniewski et al. 2007）。看起来，连接组蛋白亚型可能会有不同的蛋白质结合伙伴。例如，Reinberg 及其同事报道，人 H1b 亚型第 26 位赖氨酸的甲基化通过与恶性脑瘤蛋白 L3MBTL1 的相互作用，介导核小体阵列的高级折叠（Trojer et al. 2007）。其他连接组蛋白亚型，如 H1(0)，在第 26 位缺少赖氨酸，因此，无法与 L3MBTL1 蛋白发生相互作用（Trojer et al. 2007）。一般来说，与核心组蛋白相比较而言，连接组蛋白在物种间的保守性不是很好，因此，研究人员应该密切关注物种来源。所以，就像核心组蛋白中的情况那样，重组组蛋白 H1 亚型的使用越来越普遍。

内源性染色质结构的重建也可能需要组装含有核心组蛋白变体的核小体。例如，几乎所有的酵母 RNAPII 启动子的两侧都定位含有组蛋白 H2A 变体 Htz1（在其他物种中称为 H2A. Z 或 H2AZ）的核小体（Raisner et al. 2005；Zhang et al. 2005）。许多果蝇和哺乳动物基因似乎也有着相类似的结构（尽管不完全相同）（Mavrich et al. 2008；Schones et al. 2008）。同样，macro-H2A 变体富集在雌性哺乳动物失活的 X 染色体上（Chadwick and Willard 2001b），而 H2A-Bbd 变体在失活的 X 染色体上缺失，但可能与转录活性染色质缔合（Chadwick and Willard 2001a）。一般来说，将组蛋白变体纯化为重组蛋白，并且可从像核心组蛋白那样的未折叠蛋白质组装八聚体。但是，H2A-Bbd

变体在体外未组装成稳定的组蛋白八聚体（Bao et al. 2004）。因此，在这种情况下，H2A/H2B 二聚体和 H3/H4 四聚体必须各自组装，上面 Gloss 及其同事总结的八聚体组装方法可能最合适。

染色质重建的生化表征

或许一个研究者必须作出的最重要的决定是，用单核小体底物进行染色质分析是否充分，或者是否需要核小体阵列。如果体内研究表明一个基因调控因子作用的启动子区域或结合位点被单个定位的核小体涵盖，那么单核小体研究可能就足够了（150～250 bp）。另外，调控区可能处在定位的核小体之间，或者对单核小体组装来说太长（>250 bp）。此外，基因调控区一般在体内组织为长的核小体阵列，因此这正是研究人员期望在体外重现的生理环境。核小体阵列的体外组装经历宛如染色质压缩的复杂的染色质折叠动力学过程。因此，利用核小体阵列的研究可以让研究人员调查染色质的高级结构如何影响转录因子结合和转录活性。

如果研究者决定重建核小体阵列，那么可以组装两种一般类型的阵列。第一种是，可以利用 ACF/NAP1 组装系统来重建核小体阵列，在该系统中，核小体均匀间隔，但是就特定的 DNA 序列而言却随机定位。因此，在这种情况下，核小体不是特异地定位在基因调控序列上。结果，序列特异性 DNA 结合蛋白对调控位点的结合一般不会被核小体组装所抑制，因为这些位点会被包含在大量核小体间的连接 DNA 中（Pazin et al. 1998）。然而，这些随机定位的核小体在抑制转录起始和延伸上却非常有效，结果这类核小体阵列被广泛地用于体外转录分析（Jiang et al. 2000；An et al. 2004；An and Roeder 2004；Angelov et al. 2004；Guermah et al. 2006）。Kadonaga 及其同事已经表明，如果一个位点特异性 DNA 结合蛋白在核小体组装之前结合到 DNA 模板上，那么 ACF/NAP1 组装系统仍然可以用于产生核小体定位的局部区域。例如，他们发现细菌 LacI 蛋白的结合通过充当组装边界强行促成几个相邻核小体的定位（Pazin et al. 1997）。

第二种策略是在包含头尾相连串联重复 NPE 的 DNA 模板上，利用盐透析重建或者伴侣依赖性组装（如 NAP1）来组装核小体阵列。Simpson 等（1985）首先开发了这种模式核小体阵列模板，该模板包含 10 个或者更多的海胆 5S rDNA 序列。体外盐透析重建后，每个 5S rDNA 重复可定位一个核小体，从而产生核小体的定位阵列。这些模式核小体阵列在二价阳离子存在的情况下，在体外经历复杂的等级结构变化（Fletcher and Hansen 1996）。低浓度的 Mg^{2+} 离子（<2 mmol/L）可以通过联合相邻的核小体（“折叠”）而诱导单个的核小体阵列的分子内压缩，而 Mg^{2+} 浓度的逐渐升高（>2 mmol/L）或者低浓度的多胺（约 200 μmol/L）（Pollard et al. 1999）则可以可逆地诱导核小体阵列发生寡聚化。低 Mg^{2+} 浓度下模式阵列的分子内折叠被认为很像 30 nm染色质纤维的形成，而分子间的寡聚化产生轮廓相对清晰的可溶性结构，这些结构可以以成千上万的 Svedberg 单位沉积，并且被认为很像稳定高级染色体结构域（如染色丝纤维）的纤维-纤维相互作用（Belmont and Bruce 1994）。

在过去的 10 年中，几个研究组已创建了修改的 5S 阵列模板，该模板使转录因子结合和转录调控机制的分析更加便利。Workman 及其同事将一个包含 5 个 Gal4 结合位点

的单核小体长度的 DNA 片段插入 5S 阵列的中间，从而创造了修改的 5S 阵列。据此，他们就能够研究 Gal4 转录激活因子与核小体位点的结合（Owen-Hughes and Workman 1996）。Roeder 及其同事也创建了相似的 5S 模板，以便监测其他位点特异性转录因子的核小体结合，以及这些因子从染色质模板上驱动的转录（Malik et al. 2002；Wallberg et al. 2002）。随后，将一个涵盖腺病毒 E4 启动子的二核小体长度的片段（Utley et al. 1998）或 HIV 启动子（Steger and Workman 1997）插入到 5S 阵列中，而 Zaret 和同事则成功地插入了包含白蛋白启动子的三核小体长度的 DNA（Cirillo et al. 2002）。在所有这些案例中，核小体在周围 5S NPE 上的定位导致核小体定位在插入 DNA 片段上（见专题 13.1）。由于这些核小体阵列仅从核心组蛋白和 DNA 组装而成，它们非常适合于广泛的生物物理分析，以及体外转录分析。

一旦确定了是否重建单核小体或核小体阵列，并且选好了组装核小体的方法，研究人员在进行所期望的实验（如体外转录和染色质重塑）之前，必须花费大量的时间和精力去分析重建的染色质的质量。特别是，研究人员必须确保：①组装了真正的核小体；②按预期或期望定位了核小体；③已取得目标数目的组装核小体。

单核小体和染色小体的分析

已经组装了真正的核小体或者染色小体了吗？单核小体的重建通常通过盐透析、盐跳或八聚体转移（见上文）进行，并且涉及长为 150～250 bp 的 DNA 片段。在所有案例中，分析的第一步是在 TBE 聚丙烯酰胺凝胶（6%丙烯酰胺，丙烯酰胺与双丙烯酰胺的比例为 30∶0.8）上电泳组装反应。单核小体的迁移远远慢于游离 DNA，而且它产生一个已明确界定的蛋白质-DNA 复合物，该复合物迁移在 300～400 bp 标志 DNA 片段附近。增加组蛋白八聚体（或核小体供体）与 DNA 的比率，残留的游离 DNA 的量可减少。值得注意的是，高水平的组蛋白八聚体会造成 DNA 上的无关沉积，而这些产物在自然凝胶电泳检测中表现为缓慢迁移的拖尾。如果重建反应中包含连接组蛋白，则组装的染色小体（chromatosome）（核小体核心加一分子连接组蛋白）将比单核小体颗粒迁移得更慢。就像在核心组蛋白的情况下那样，应时常在连接组蛋白与核心组蛋白的不同比例下进行多重组装反应，并且应通过自然凝胶监测重建的效率。即使在染色小体形成后，连接组蛋白也会继续结合，从而产生在自然凝胶上缓慢迁移的异常复合物。

如果自然凝胶分析表明单核小体或染色小体已经被组装，那么有两个简单的方法可用于确保核小体有正常的组蛋白化学计量和正确的 DNA 含量。为了评估组蛋白的化学计量，必须首先通过 5%～30%的甘油梯度沉降纯化单核小体或染色小体，以去除未组装的组蛋白八聚体或核小体供体。之后，峰馏分用 18%的 SDS-聚丙烯酰胺凝胶电泳（SDS-PAGE）分析，紧接着考马斯亮蓝染色显示组蛋白。一般来说，组蛋白会被多数银染方法拙劣地且异质地染色，这就是为什么通常尽量避免使用这些方法的原因。大多数核心组蛋白迁移成一组 4 个紧密间隔的多肽，迁移的顺序为：H3、H2B、H2A 和 H4（从最慢到最快）。若是酵母组蛋白，H2B 和 H2A 会作为一个单一种类迁移。正确组装的核小体含有相同水平的所有 4 种组蛋白。

为了评估 DNA 含量，样品用微球菌核酸酶广泛消化，产物在高百分比的琼脂糖凝胶或变性 PAGE 中解析。一个典型的单核小体可以从广泛的 MNase 消化中保护 147 bp

的 DNA，但最终核小体 DNA 也将被消化。同样，在染色小体的情况下，在酶抵抗性更强的 147 bp 产物积累之前的消化时程中，MNase 消化产生一个瞬时的约 166 bp 的“染色小体停留（chromatosome stop)”。

*核小体定位的分析：单核小体。*当试图重建一种内源性染色质结构或进行染色质的机制性研究时，组装相对于相关 DNA 序列精确定位的单核小体往往是至关重要的。这种“平移”定位可以确保目的 DNA 结合蛋白的识别序列是否被缠在组蛋白八聚体上。除了平移定位，DNA 也可以在组蛋白八聚体表面采取特定的旋转位置。旋转定位的核小体可以有一个或几个不同的平移位置，但群体中的所有分子在八聚体上都有固定朝向的 DNA 螺旋，也就是说，DNA 的特定片段总是背对八聚体表面，而 DNA 的另一段，可能总是向内面对八聚体。

生理相关的核小体位置有时可用自然基因序列重建（Roberts et al. 1995；Terrell et al. 2002)，但是在很多情况下，相对于其他靶序列，良好界定的异源性 NPE 被用于定位核小体。NPE 是高亲和性组蛋白八聚体结合位点，它的碱基组成有利于 DNA 在八聚体周围的弯曲（Lowary and Widom 1998)。最早鉴定的 NPE 之一是 Shrader-Crothers序列元件，它由 4～10 个富含 A/T 的 DNA 序列串联拷贝组成，其一致序列为：$(A/T)_3$NN $(G/C)_3$NN（Shrader and Crothers 1989)。改变 NPE 相对于目标 DNA 序列的位置，可让研究者确定目标 DNA 序列的方向是“面对”还是“背对”八聚体表面。例如，Imbalzano 等（1994）使用多拷贝的 Shrader-Crothers 序列确定 RNAPII 启动子的 TATA 盒在核小体表面的方向。此外，通过改变定位序列和 TATA 盒的间隔，这些作者能定位这些启动子元件是背向还是面向八聚体表面。

两个应用更广泛的 NPE 是来自海胆（Simpson et al. 1985；Hansen et al. 1989；Dong et al. 1990）或爪蟾（*Xenopus borealis*）的 5S rRNA 基因（Hayes et al. 1990)，以及被称为 601 的合成 DNA 序列（Lowary and Widom1998)。601 NPE 是独一无二的，因为它能精确定位核小体，即使 601 序列被嵌入一个大的 DNA 片段中也是如此(Shundrovsky et al. 2006)。相比之下，5S rRNA NPE 在一个主要位置定位约 50%的八聚体，其余 50%的分子定位于次要位置，这些次要位置与主要位置差 10 bp 的整数倍(Dong et al. 1990)。这种核小体定位的异质性常常在自然 PAGE 分析重建单核小体期间表现得很明显。例如，当在 5%的自然 PAGE 胶中分析 5S 单核小体组装时，不同的平移位置清楚地显示为一系列不同迁移率的条带——定位在片段中心的核小体迁移得最慢，而靠近 DNA 末端的核小体则迁移较快。相反，601 单核小体在 PAGE 分析中会产生明确界定的单一类型的条带。一般来说，当研究者希望将启动子或者转录因子结合位点放置在定位的核小体附近时，601 和 5S NPE 都会表现出最大的效用。到目前为止，很少尝试在 601 或者 5S 中插入结合位点或启动子，尽管在一个案例中，Pederson 及其同事成功地将酵母热激因子（HSF）的结合位点植入 5S NPE 中，却没有破坏核小体定位（Pederson and Fidrych 1994)。

核小体的位置可以通过不同核酸酶的消化进行作图（见第 11 章)。首先，研究人员可以用一组限制酶提供核小体定位的定性评估。核小体组装抑制限制酶切割效率达 10^3～10^5 倍，与更加周边的地点相比，在核小体二分体上有着更强的抑制作用。对切

割底物的定量分析可简单估算核小体的密度和定位。如果将单核小体首先用 MNase 修剪到 147 bp 的核心粒子，限制酶消化可提供更多的定位信息，然后在纯化的 DNA 上进行限制性内切核酸酶消化。在自然聚丙烯酰胺凝胶上电泳少量消化产物，确定相对于限制性酶切位点的核小体边缘。值得注意的是，此方法要求均匀标记的 DNA 片段，以用于 Southern 杂交分析，或足够数量的重建产物，以便用溴化乙锭检测。

在单碱基对分辨率水平上进行核小体位置作图的一个常见方法是通过外切核酸酶 III 消化（见图 13.2）。如在第 11 章中的详细描述，*Exo*III 是进行性 3′—5′的外切核酸酶，它的活性在遇到 DNA 结合蛋白（如核小体）时受阻。*Exo*III 反应产物经高分辨率变性聚丙烯酰胺凝胶电泳后，可以根据 5′^{32}P 标记确定核小体一个边缘的位置。为了精确确定核小体的平移位置，必须进行两个不同的核小体制剂的 *Exo*III 分析，每个制剂使用不同末端^{32}P 标记的 DNA 探针。

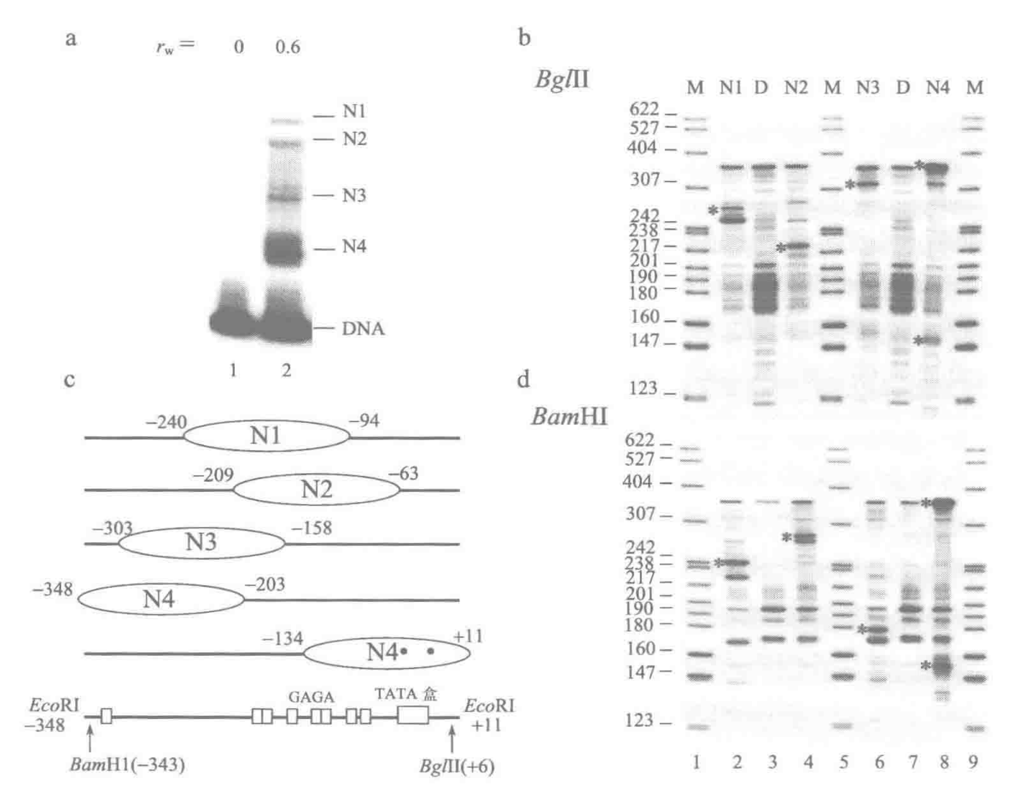


图 13.2 用自然凝胶电泳和 *Exo*III 分析探测单核小体的平移定位。(a) 自然聚丙烯酰胺凝胶电泳检测以重量比（r_w）为 0.6 的组蛋白∶DNA 重建的单核小体；DNA 为含有 *hsp*70 启动子的放射性标记 359 bp *Eco*RI 片段。4 种主要的核小体种类显示为 N1、N2、N3 和 N4。(b) 从凝胶中切出对应于每一个核小体种类的条带（在 *Bgl*II 或 *Bam*HI 位置被放射性的 DNA），再洗脱核小体，最后用 400 U/ml 的 *Exo*III 37℃消化 2 min。DNA 片段在测序胶上的电泳分析，泳道代表 N1～N4 核小体；(d) 游离 DNA；(M) pBR322 *Hpa*II 消化分子质量标准物；（星号）核小体边界。(c) 定位在 359 bp *Eco*RI 片段（线条）上的核小体用（椭圆）表示，图中给出±2 bp 的核小体边界位置。图中还标出了 GAGA 和 TATA 启动子元件（经 Elsevier 许可转载自 Hamiche et al. 1999，© 1999）。

由于 DNA 缠绕在组蛋白八聚体周围，小沟以 10 bp 周期性地被暴露在核小体的表面。如果 DNA 是旋转定位在八聚体表面，在小沟切割 DNA 的试剂中会产生 10 bp 的高敏感位点梯状条带。通常，使用典型的指纹反应，用 DNase I 或羟基自由基（第 11 章）定位核小体 DNA 的旋转档位（rotational setting）；指纹反应利用的是末端标记核小体的有限消化（图 13.3）。通过这种方式，可以确定因子结合位点是否被暴露在核小体上或者位点是否并置在八聚体表面。这种分析对研究 ATP 依赖性染色质重塑酶 SWI/SNF 家族如何破坏核小体结构也是有帮助的（Cote et al. 1994；Imbalzano et al. 1996；见下文）。

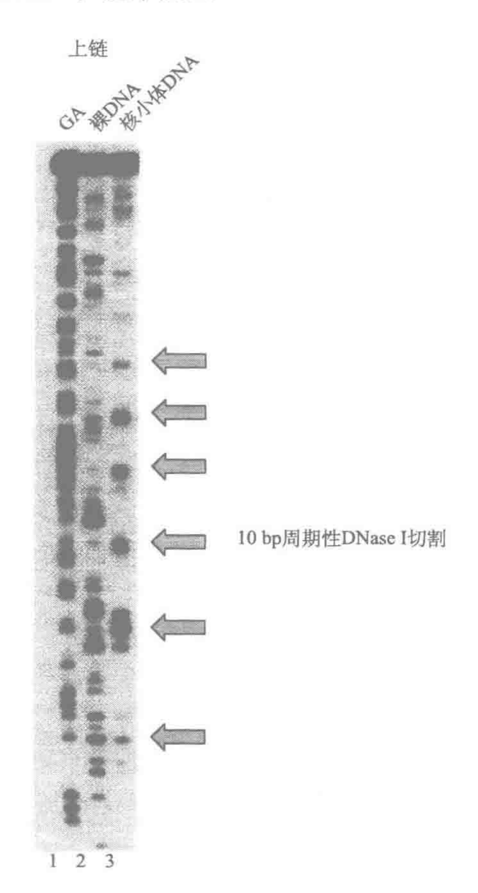


图 13.3 通过 DNase I 消化探测核小体 DNA 的旋转定位。DNase I 消化用于评估重建单核小体上的 DNA 旋转定位。在这个实验中，通过盐透析或八聚体转移法将单核小体组装到一个^{32}P 末端标记的 DNA 片段上。用 DNase I 有限消化重建的单核小体，并纯化 DNA 产物，然后在变性聚丙烯酰胺凝胶中分离。如果 DNA 旋转定位在组蛋白八聚体表面，这种 DNase I 消化将会产生一个 10 bp 切割和保护的诊断模式，指示 DNA 小沟周期性暴露在核小体的表面。值得注意的是，DNase I 呈现相当大的序列偏好，实验中应始终平行分析游离 DNA 的 DNase I 消化产物（经 MacMillan Publishing Ltd. 的许可修改自 Imbalzano et al. 1994，© 1994）。

核小体阵列的分析

核小体定位。一般通过以下两种方法之一完成线性或环状核小体阵列的重建：盐透析或 ATP 依赖性 ACF/NAP1 系统。盐透析重建是最适合包含首尾相接核小体定位序列重复 DNA 模板的方法，因为这个方法在异源 DNA 片段上产生紧密排列的非生理间距核小体（见专题 13.1）。ACF/NAP1 重建可以在线性或者环状 DNA 上产生均匀间距的核小体，但需要注意的是重建产物可能需要进一步纯化以除去组装因子。在这两种情况下，通常要通过限制酶或 MNase 消化监测核小体定位。阵列 DNA 对限制酶消化的

可及性能探测阵列内一个或多个核小体，而有限的 MNase 消化则能提供关于核小体间隔的信息。如果重建产生一个定位核小体阵列，那么有限 MNase 消化将产生一个重复“梯子”的 MNase 消化产物，该产物反映了每个暴露的连接 DNA 内的切割（专题 13.2 和专题 9.2）。“梯级”间的距离显示了定位核小体之间的间距，并且梯子上“梯级”的个数反映了定位阵列的质量（专题 13.2）。例如，在外源 DNA 片段上盐透析重建可能只产生 1～3 个条带，因为这样的重建含有密集排列的核小体斑块及无核小体的间隙。然而，在相同片段上的 ACF/NAP1 组装可能产生 10 多条条带。通过改变 ACF/NAP1 组装反应的盐浓度，也可以调整核小体之间的间距（Blank and Becker 1995）。这些相同的 MNase 消化也能用于确定阵列内核小体的定位（专题 13.2）。如果阵列是在末端标记的 DNA 片段上重建，那么 MNase 梯子的梯级将提供核小体定位的低分辨率视图。另外，MNase 消化产物的 Southern 印迹分析，利用寡核苷酸作探针，可用于确定某种特定序列是否包围在核小体中（专题 13.2；也见 Sinha and Peterson 2008）。也可以利用连接酶介导的聚合酶链反应（LM-PCR）或者循环引物延伸进行更高分辨率的 MNase 切点作图。例如，Laybourn 及其同事使用引物延伸分析进行单核苷酸水平上的 MNase 切点作图，以便在一个长 DNA 片段上完成 NAP1-依赖的核小体组装之后，追踪酵母 PHO5 启动子上的核小体定位（Terrell et al. 2002）。

专题 13.2 核小体阵列的 MNase 分析

微球菌核酸酶优先切割核小体之间的连接 DNA 中的核小体 DNA，这里展示了三种不同的使用 MNase 探测重建阵列内核小体定位的方法。在图 13.4a 中，阵列用浓度（或消化时间）逐渐增加的 MNase 处理，脱蛋白质后的样品在琼脂糖凝胶上进行电泳。用溴化乙锭染色凝胶显示 DNA，出现相隔约 150 bp 的一系列规则条带，表明核小体在阵列内是均匀间隔的。与所示例子相反的是，未显示均匀间隔的重建其特点是出现单核小体条带（迁移最快的约 150 bp 的条带），并且较大的 N-mer 产物从几乎没有到完全没有。图 13.4b 说明了一个测试阵列内核小体是否平移定位的间接末端标记实验。在这种情况下，阵列重建在末端标记的 DNA 上，或用末端定位探针（阵列上以小线段显示）进行 Southern 分析，以用于核小体位置作图。阵列经 MNase 部分消化，确定相对于直接或间接标记末端的 MNase 切割的位置。一系列受保护区域指示出了定位的核小体，而均匀间隔的阵列（而不是随机定位的阵列）将产生 MNase 切割的游离 DNA 模式。需要注意的是，如果使用末端标记阵列，高浓度的 MNase 将切除标记端，导致信号丢失。图 13.4c 用于确定一个小的目标序列（小于 100 bp）是否位于阵列群体内的核小体中。在这种情况下，按图 13.4a 中的方法进行 MNase 消化，用探测目标 DNA 的寡核苷酸进行 Southern 印迹分析。如果目标 DNA 被群体中的绝大多数阵列中的核小体包围（探针 1），则寡核苷酸与单核小体 DNA 和 MNase 产物的梯形条带杂交（左）。但是，如果目标 DNA 主要位于核小体之间的连接 DNA 上（探针 2），则寡核苷酸将检测小于 150 bp 的 DNA 片段，并且核小体梯形条带会不太明显。值得注意的是，如果核小体在群体中是随机定位的，那么结果将是所显示的两种极端情况的组合。

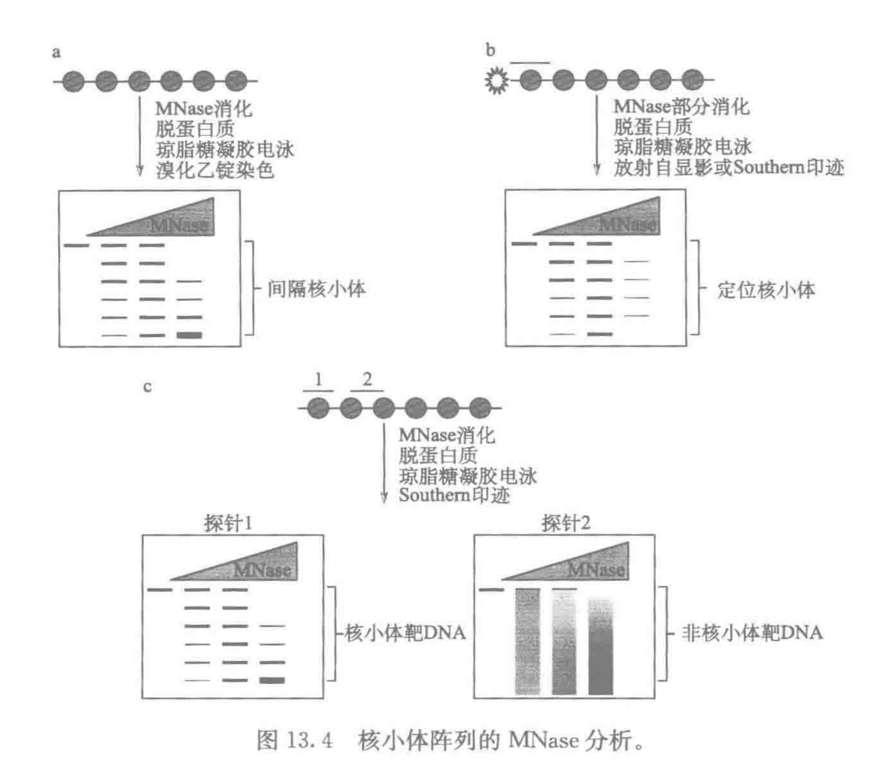


图 13.4　核小体阵列的 MNase 分析。

核小体密度分析。当组装核小体阵列时，最关键的是确定该模板 DNA 被核小体饱和的程度。在 12 个核小体的线性阵列上，即使是单个无核小体的间隙，也都可以阻止这些阵列折叠成 30 nm 样纤维（Hansen and Lohr 1993）。未饱和也能产生核小体模板，只不过在这种情况下群体中的许多分子上，启动子元件是无核小体的。此外，如果组装反应中加入太多八聚体，或者 DNA 的量太低，阵列可能会过饱和，以致过量的八聚体非特异性地结合核小体 DNA 并干扰染色质折叠或随后的功能分析。

核小体的饱和度评价通常涉及多种技术，但总体策略取决于重建方法。当阵列组装在缺乏重复核小体定位序列的 DNA 模板上时，核小体的饱和程度将通过 MNase 的消化和拓扑分析来确定。当这种阵列是通过 ATP 依赖性组装反应重建的，有限的 MNase 消化对于线性和环状 DNA 模板都是有用的，并且消化梯子中“梯级”的数目为重建核小体的平均数目提供了一个指标。对小于 147 bp 的 MNase 产物的检测表明，存在大量无核小体的 DNA。这种分析对于盐透析或分子伴侣介导的组装没有太多益处，因为重建核小体的非均一性间距会破坏 MNase 梯状条带。重要的是，MNase 活性似乎没有受到非特异性结合的组蛋白八聚体的抑制，因此 MNase 消化可误导对组装反应质量的看法。也就是说，即使获得了 MNase 产物的规则的梯状条带，样品仍有可能以很高的水平非特异性结合组蛋白八聚体，它们可能会干扰后续的分析。

评估环状 DNA 模板上核小体饱和度的最好的办法是拓扑分析。组蛋白八聚体周围 DNA 的缠绕将 DNA 限制为负超螺旋，同时在相邻的连接 DNA 中产生正超螺旋。如果

组装反应中出现真核 Topo I，那么核小体的饱和程度可通过监测被引入封闭环状质粒的负超螺旋的个数直接进行评估。一般情况下，研究人员会以经过 Topo I 松弛化预处理的质粒 DNA 开始核小体的重建，以确保随后检测到的任何超螺旋都是由核小体组装产生的。一维凝胶分析可以近似地测量核小体组装程度，但二维分析可以更准确地评估核小体密度；在二维分析中，样本是在有或没有 DNA 插入药物氯喹的情况下电泳的（详细方案见 Fyodorov and Kadonaga 2003）。这种拓扑分析最好是与 MNase 分析一起进行，因为 DNA 在多种 DNA-结合蛋白周围的缠绕能改变这种分析中质粒的拓扑结构。

包含头尾相连 NPE 重复的 DNA 模板已经被工程化，以便将限制性酶切位点置于每个定位核小体之间的连接 DNA 上。例如，5S rDNA 阵列在每个 5S 重复之间都包含 *Eco*RI 位点，而 601 序列阵列在重复间包含 *Sca*I 位点。当将这些阵列模板重建成核小体阵列时，可以很容易地通过一个简单的限制酶消化及后续的凝胶电泳确定核小体的饱和程度（专题 13.3）。本质上，这种分析非常类似于如何检测单核小体的组装。例如，当一个重建阵列只包含少数核小体（未饱和）时，相对于释放的单核小体，限制消化或凝胶分析会显示出大量的释放的游离 DNA 重复。相比之下，一个核小体完全饱和的重建阵列通过这种分析大多会产生单核小体。通常，我们发现包含约 4%游离 DNA 重复的重建阵列代表了一个“完全”饱和的核小体阵列（注意组蛋白八聚体周围 DNA 的缠绕会淬灭约 2.5 倍的溴化乙锭荧光）。企图完全消除游离 DNA 通常导致含有过量非特异性结合八聚体的阵列，表现为泳道中消化产物的普遍拖尾。

专题 13.3　核小体阵列的 *Eco*RI 酶切分析

由头尾相接 NPE 重复构成的模式核小体阵列模板，在每个 NPE 之间都含有限制酶识别位点（图 13.5，左图）。含有 5S rDNA 基因的模板在每个重复之间都有 *Eco*RI 位点，而含有 601 NPE 的模板则包含 *Sca*I 位点。盐透析重建后，核小体阵列用 *Eco*RI（或 *Sca*I）消化，以确定被核小体占据的 NPE 的百分比。消化产物在 5%聚丙烯酰胺凝胶中电泳，再用溴化乙锭染色凝胶（图 13.5，右图）。在较低的组蛋白与 NPE 比值（r）（$r = 0.6$）下，很多 NPE 无核小体（图 13.5，右图）。当 r 值增大时，大部分 NPE 组装成单核小体。应该注意的是，*Eco*RI 确实产生一些二聚体和三聚体产物，因为存在阻止 *Eco*RI 到达部分

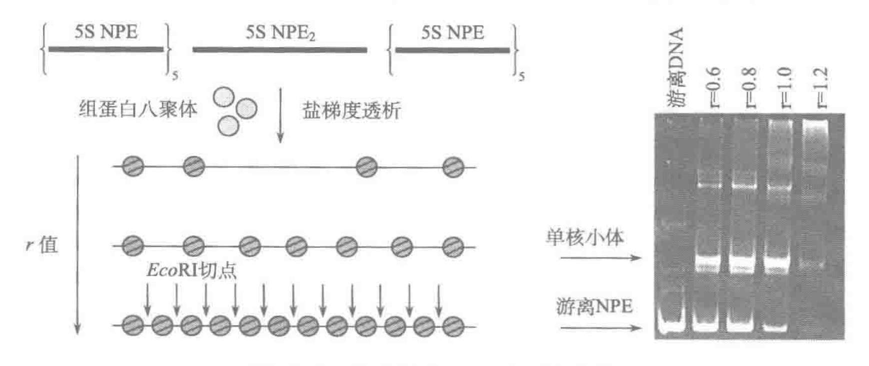


图 13.5　核小体阵列的 *Eco*RI 分析。

NPE的另类核小体定位。以核小体完全饱和的重建通过*Eco*RI分析会包含约4%的游离DNA（注意：单核小体中的DNA少结合溴化乙锭约2.5倍）。在这组重建中，$r=1.0$的样品非常接近饱和阵列。在较高的r值，额外的八聚体非特异性结合核小体阵列，在分析中产生迁移更缓慢的拖尾（如$r=1.2$的重建）。

用于分析重建核小体阵列质量的最严格的方法是在分析超速离心机中进行速度沉降实验。这是一个过时的技术，但随着最先进的分析超速离心机（如Beckman XL-A和XL-I）的出现，它正在卷土重来（综述见Hansen et al. 1994，1997；Ausio 2000）。分析超速离心机方案让研究人员可确定样本的纯度、分子质量、沉降系数、扩散系数、摩擦系数（形状），以及结合/解离反应。的确，这种分析方法对任何大分子的溶液态行为的深入分析是必不可少的。一般情况下，我们用速度沉降来精确确定核小体的饱和度、鉴定劣质的阵列和（或）同质的阵列，并表征由缓冲液条件或由非组蛋白诱发的阵列的高级折叠。例如，当208-12 5S阵列（12个重复）被核小体饱和时，在低盐TE缓冲液中，该阵列在XL-I超速离心机中会沉降为均匀的29S大分子群，但只包含11个核小体的阵列会沉降在27S处，其他方法都不能达到这一灵敏度。此外，当核小体阵列折叠成高级结构时，这会导致形状上的变化，这种形状变化改变摩擦系数，从而改变沉积速率。的确，对于溶液中染色质折叠动力学的分析，沉降速度分析法是为数不多的现成的方法之一（Hansen et al. 1989；Hansen and Lohr 1993；Hansen and Wolffe 1994；Carruthers and Hansen 2000；Horn et al. 2002a；Dorigo et al. 2003；Shogren-Knaak et al. 2006）。虽然我们日常能在分析超速离心机中分析阵列，但大多数研究人员并没有获得这样的设备。因此，简单的*Eco*RI或*Sca*I分析仍然是分析在5S或601重复序列上组装的线性核小体阵列的一个最好的方法。

体外测验组蛋白修饰作用策略

早在1964，Allfrey等（1964）就建议在组蛋白氨基末端结构域的可逆性乙酰化与转录之间建立一种联系。许多支持组蛋白高度乙酰化（尤其是H3和H4）和转录活跃染色质之间存在相关性的研究都遵照了这一提议。这些早期的研究鉴定了核心组蛋白上的几个其他翻译后修饰，但过去10年的工作已经阐明了大量的核心组蛋白修饰，包括丝氨酸或苏氨酸磷酸化、精氨酸甲基化、脯氨酸异构化，以及赖氨酸的甲基化、泛素化、乙酰化或SUMO化（最近的一篇综述，见Kouzarides，2007）。事实上，位点特异性组蛋白修饰的调控被认为是染色质调控转录的主要手段之一。

正如在第一章中所概括的，各种模式的组蛋白修饰与不同的转录状态相关。由Pol II转录的基因的普遍特征包括H3中几个赖氨酸（主要是K9、14）的乙酰化；H4K4、K8、K12和K16的乙酰化；H3K79的甲基化，以及H3K4的双甲基化或三甲基化。某些转录基因可能也可以通过H3K14乙酰化和H3S10磷酸化加以区别。相比之下，被抑制的基因座位趋向于缺乏这些标记，并且取而代之的是富集含有几个甲基化赖氨酸的组蛋白，如H3K9me2、H3K27me2或H4K20me3。但是，似乎并不存在活跃转录状态或抑制转录状态本身的“密码”，并且研究者最钟情的受调控基因可能有其自己的修饰模

式。至于修饰模式是如何被“读”入转录决策的，则是通过生化方法可以解决的一类问题。在这方面，对于染色质研究团体的一个主要挑战在于制备纯化的组蛋白，并且这些纯化的组蛋白含有均一的、明确的翻译后修饰模式，以便研究转录激活或抑制的生化机理。下面的几节将考虑三个策略，包括：①组蛋白修饰的酶加入；②自然化学连接；③甲基赖氨酸类似物的引入。被修饰组蛋白的应用将在后面的一节中加以讨论。

组蛋白修饰的酶加入

自1996年Allis及其同事（Brownell et al. 1996）发现第一个核组蛋白乙酰转移酶Gcn5以来，已经鉴定和生化表征了大量的组蛋白修饰酶。在许多情况下，重组酶能很好地作用于组蛋白底物（如Gcn5、P/CAF、CBP、hMOF、R-SET7和G9A），并且有些研究利用这些酶创建了含有位点特异性乙酰化或甲基化组蛋白的核小体或核小体阵列，以便用于功能分析（Hassan et al. 2001；An et al. 2004；Chandy et al. 2006；Trojer et al. 2007）（也见下面图13.7）。例如，Workman及其同事用几个酵母组蛋白乙酰转移酶（HAT）复合物对核小体阵列进行乙酰化，并监测对HIV-1转录的影响（Steger et al. 1998）。同样地，Roeder及其同事在固定模板分析中使用P300 HAT确定转录激活期间哪些步骤是由组蛋白乙酰化刺激的（An et al. 2004）。原则上，组蛋白的酶修饰可以用游离组蛋白或核小体作为底物，但是在这两种情况下，必须滴定酶量及反应时间，以限制次要修饰产物。事实上，这种方法的主要限制之一是，酶修饰可引起位点特异性的异质性。酶加入策略的另外一个缺点是，非常难于驱使这些类型的反应完成。此外，所需的组蛋白修饰酶可能根本就得不到，或者难以纯化到足够的量。此外，如果一个研究者希望创建一个复杂的组蛋白修饰模式，那么这种策略可能行不通，因为这将需要多重酶加入。因此，虽然这一策略可以在某些应用中非常有用，但局限性还是很多。

自然化学连接

在过去几年中，自然化学连接已经成为创造氨基末端修饰的重组组蛋白的一个有力工具（He et al. 2003；ShogrenKnaak et al. 2003，2006；Shogren-Knaak and Peterson 2004；Ferreira et al. 2007）。原理上，这种方法非常适合创建具有多种类型修饰残基和修饰残基组合的重组组蛋白。此外，不像酶修饰组蛋白或细胞裂解物中纯化的组蛋白，自然化学连接提供均匀修饰的组蛋白。这种自然化学连接策略从未保护的合成肽和表达的蛋白质片段产生全长蛋白质，这对纳入通过遗传密码不能被直接掺入蛋白质的氨基酸非常适合。具体而言，这一技术需要一个以羧基末端硫酯结尾的氨基末端的多肽片段，以及一个以氨基末端半胱氨酸开头的羧基末端多肽片段（图13.6）。当混合在一起时，这些片段可以通过羧基末端硫酯和氨基末端半胱氨酸巯基的反式硫酯化形成进行可逆共价缔结。这一反应中间物然后可以通过*S*-酰基到*N*-酰基的转移发生不可逆的重排，导致在连接位点处出现一个经典的酰胺肽键。此方法与其他含有半胱氨酸的蛋白质兼容，因为具有这些半胱氨酸的硫酯产物不能重排形成酰胺键，且硫酯化是可逆的。

为了将感兴趣的修饰残基掺入被连接的组蛋白的氨基末端结构域，可以通过固相肽合成产生含有氨基末端硫酯的片段，并且从细菌中表达和纯化含有羧基末端半胱氨酸的蛋白质片段。重要的是，我们发现起始甲硫氨酸在细菌中被有效地从包含相邻半胱氨酸残基的组蛋白中去除，这就不再需要表达含有位点特异性蛋白酶裂解位点的组蛋白。从

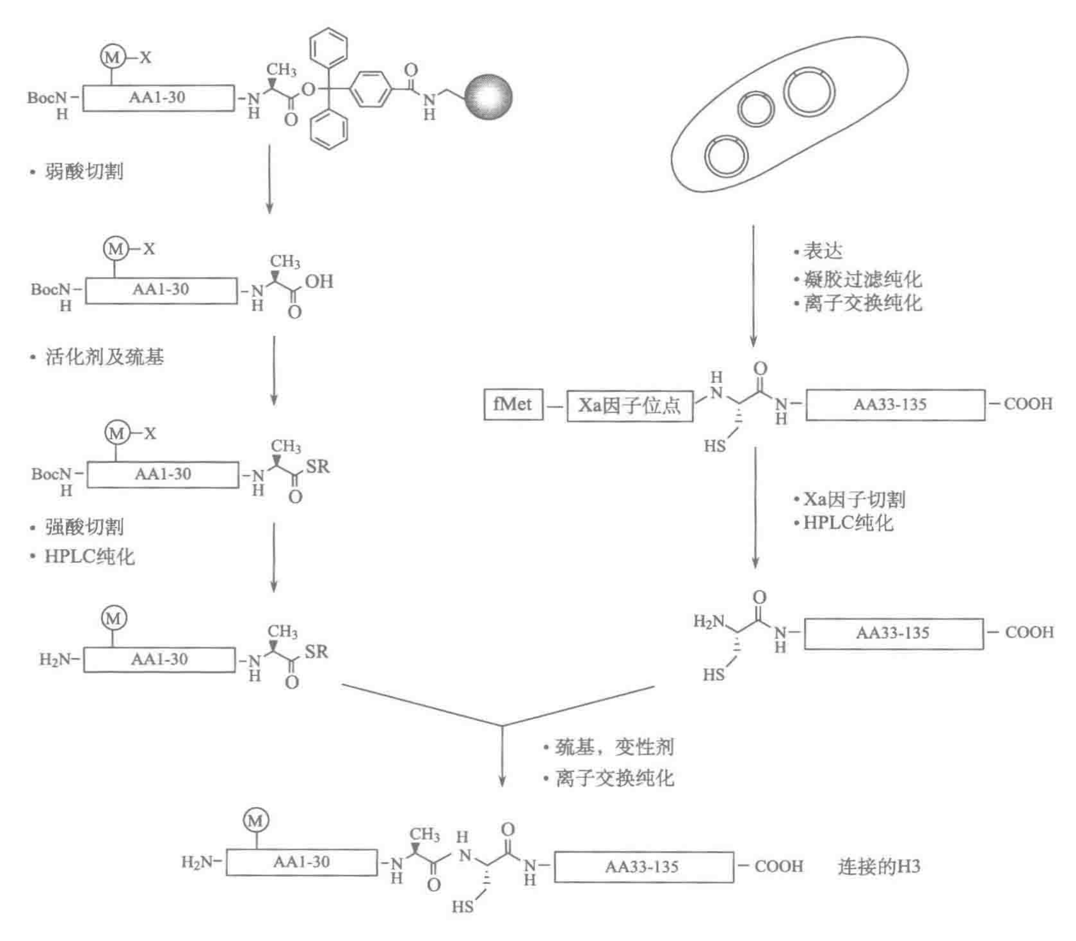


图 13.6　产生“设计师”组蛋白的自然化学连接策略。在示例中，通过标准的固相肽合成方法在酸过敏性支持物（左）上生产一个组蛋白 H3 的氨基末端肽片段，该片段包含一个特定修饰的氨基酸残基（带圈的 M 标示的残基）和一个羧基末端硫酯基团（COSR）。图中还显示了一个从细菌中表达和纯化的含有氨基末端半胱氨酸残基的组蛋白 H3 羧基端蛋白质片段（右）。如图中显示的那样，可通过蛋白质水解切割暴露出氨基末端的半胱氨酸，或毗邻起始甲硫氨酸残基设置半胱氨酸，起始甲硫氨酸可以在细菌细胞内被有效切割。在巯基试剂的存在下，这两个片段的反应生成含有目的修饰的自然全长组蛋白 H3（经 Elsevier 许可转载自 Shogren-Knaak and Peterson，2004，© 2004）。

理论上讲，通过表达含有羧基末端硫酯的重组氨基末端片段（由内含肽切割获得；David et al. 2004），以及合成含有氨基末端半胱氨酸的羧基末端尾肽，研究羧基端组蛋白修饰也应该是可能的。

要生成具有反应活性羧基端硫酯的肽，则需要在其他潜在反应性官能团存在的情况下，末端肽 α-羧酸基团选择性地硫酯化。为了取得这一选择性，要通过标准的基于 Fmoc 固相肽合成方法在酸过敏的树脂上合成肽。我们已发现，所有残基的双偶联可提高组蛋白肽的产量和纯度。然后，用弱酸性切割混合物将合成的肽从合成支持物上切割下来，暴露游离的末端 α-羧酸，同时保持侧链保护基团。之后，用标准肽偶联试剂硫

酯化肽末端 α-5 羧酸。随后用强酸性切割混合物处理，从而去除侧链保护基团，同时保持羧基末端硫酯化。最后，用高压液相色谱法（HPLC）获取硫酯连接底物纯品。有关此方法的详细描述，见 Shogren-Knaak 和 Peterson（2004）。

连接反应的产物含有在自然组蛋白中不存在的半胱氨酸残基。当设计半胱氨酸引入的位置时，必须考虑几个标准。首先，连接必须发生在氨基端最末端的 30～35 个残基中，因为无法有效地合成更大的肽。其次，应尽量鉴定一个在真核生物间不是很保守的氨基末端残基，一种策略是在所需位置的组蛋白氨基末端结构域中插入一个新的半胱氨酸残基，但须注意这可能改变潜在功能元件的间距。最后，脯氨酸、缬氨酸或异亮氨酸已知会抑制连接，在重组蛋白中不应与半胱氨酸残基（羧基末端）相邻。

优点和缺点。自然化学连接法与其他策略相比有几个明显的优势。例如，研究人员完全控制掺入哪些修饰残基，以及它们在氨基末端结构域中的位置，产物同质性能得到保证，在单一肽中可掺入的修饰的个数不受限制。但是，这种方法也有几个缺点。首先，时间和资源投入巨大。除非研究人员有自己的肽合成设备，否则获得在非常规的酸不稳定的树脂上合成的高品质的肽很困难。特别是，我们发现很难获得超过 30 个残基的肽，因为它们往往被截断或纯度较低。其次，一旦组蛋白氨基末端肽为硫酯化反应做准备而从树脂上被切割下来，它们的带电性质可以引起溶解度问题，这大概是由保留在氨基末端和氨基酸功能基团上的疏水性保护基团的继续存在造成的。因此，大多数反应必须在有高浓度变性剂（如 6 mol/L 盐酸胍）的情况下进行。就某些甲基化 H3 肽来说，因为严重的溶解度问题，我们一直无法进行连接反应。另外一个问题是，由于肽合成的限制，这种方法对发生在羧基末端约 30 个残基中的修饰不适合。不过，考虑到连接策略的强大可能性，至少这些问题中的某些问题有望通过商业化肽合成设施得到解决和尽量减少。

甲基赖氨酸类似物引入

在过去几年中，已经很清楚，组蛋白修饰并不限于核心组蛋白的氨基末端结构域，似乎有许多修饰位于羧基末端的球状结构域中。在这些标记中，比较显著的标记之一是 H3K79me3，在散装细胞染色质中，它似乎是一个丰富的标记，并且在修复 DNA 双链断裂期间，可能在 DNA 修复蛋白的招募中发挥关键作用（Huyen et al. 2004）。已经通过细胞组蛋白的质谱分析鉴定了许多其他羧基末端修饰（Zhang et al. 2003；Garcia et al. 2007），我们期待着它们的功能分析。

如上所述，自然化学连接策略对发生在羧基末端约 30 个残基的修饰是不适用的，同时对于没有修饰酶的新标记也是不适用的。研究人员如何能在体外研究这些新标记呢？最近，Shokat 及其同事通过利用一个旧化学位并将其应用于重组组蛋白生物化学，已经解决了这个问题，至少对甲基化赖氨酸残基来说是如此。在他们的甲基类似物方法中，使用氨乙基化反应将半胱氨酸转换成一个 *N*-甲基化氨乙基半胱氨酸残基，这是一个甲基赖氨酸类似物（MLA）（Simon et al. 2007）。值得注意的是，甲基化赖氨酸和 MLA 之间的唯一差别是赖氨酸的 γ 亚甲基被硫化物取代，这几乎不产生结构变化。这个策略利用一个事实，即组蛋白只含有一个半胱氨酸残基，即 H3C110，它可以用丙氨酸替换，而不会影响核小体结构。然后，可以用一个独特的半胱氨酸残基替换感兴趣的赖氨酸，之后可将这个半胱氨酸残作为氨乙基化的靶标。重要的是，MLA 技术允许所

有三种甲基化状态的特异合成，并将其作为被修饰组蛋白分子内的一个同质群体。因此，使用特定的反应条件和商业化获得的（2-卤乙基）胺，可将毫克量级的组蛋白转换为 1-me MLA、2-me MLA 或 3-me MLA，且产量超过 90%、成本低。一般情况下，所需要的试剂的数量和专业知识似乎对大多数分子生物学或生物化学研究组是合理的。如需详细方案，见 Simon 等（2007）。

已经有一些前期研究的报道，其中已将 MLA 用于分析甲基化组蛋白残基在染色质结构或功能中的作用。最近，Luger 和同事们已经利用这种策略来探讨含有 H4K20me3 或 H3K79me2 的核小体和核小体阵列的结构特性（J. Hansen，个人通讯）。他们发现，H4K20me3 增强了核小体阵列的盐依赖性压缩，而 H3K79me2 改变了单个核小体表面的电荷特性，后者可能会影响非组蛋白的结合。在原始报告中，Shokat 及其同事创建了含有 H3K9me2、H4K20me2 和 H3K79me2 的重组组蛋白（Simon et al. 2007）。令人惊讶的是，在这些 MLA 中，每一个都对相应的甲基赖氨酸特异性抗体具有活性，H3Kc9-2me 能特异性地与组蛋白结合蛋白 HP1 相互作用。MLA 没有显示出严重破坏核小体结构，因为在 H3K9 上带有氨乙基半胱氨酸的单核小体仍然可以被 SUV39H1 甲基化，而且这些带有 H3K9、H3K20 或 H3K79 MLA 的单核小体可以被 ACF 重塑复合物调动起来。

优点和缺点。MLA 策略目前是专注于位点特异性组蛋白甲基化研究的首选方法。化学方法似乎很强大，而且试剂便宜，同时对 MLA 引入重组组蛋白的位置没有限制。本方法为数不多的局限性之一是，研究人员对每个组蛋白只能引入一种类型的甲基化状态（即同一个组蛋白上引入一个或多个 Kme3，而不能同时引入 Kme2 和 Kme3)，虽然多个甲基化状态可被引入到不同的组蛋白上（即 H3Kme2 和 H4Kme3）。因此，研究人员可以产生许多模式的甲基化赖氨酸，但有一些限制。此外，通过将 MLA 方法与自然化学连接相结合，研究人员可以将位点特定性赖氨酸甲基化与其他组蛋白标记结合起来。但是，我们注意到，在研究者中仍然有一些顾虑，那就是，MLA 可能不能忠实地模拟甲基化的赖氨酸，而且额外的硫化物基团可能干扰或减少蛋白质-组蛋白相互作用。随着越来越多的研究人员使用这种技术，有可能阐明 MLA 的其他优点和缺点。

染色质重塑/修饰酶的分析

当早期研究揭示真核细胞核内的 DNA 是通过与组蛋白的缔合而被压缩时，研究人员立即意识到这种压缩结构将对基因表达的调控产生严重的影响。研究人员发现，活性座位呈现对核酸酶敏感性的普遍增加，这一发现迅速导致染色质结构可能通过酶活性调控的观点，因此转录也同样可能通过酶活性调控（Weintraub and Groudine 1976）。虽然对这种酶的寻找开始于 20 世纪 80 年代初，但直到 1994 年才从酵母和人类细胞中鉴定出第一种“染色质重塑”酶，这就是 ATP 依赖性重塑酶 SWI/SNF（Cote et al. 1994；Kwon etal. 1994）。这一发现之后又鉴定了很多 ATP 依赖性重塑酶相关家族。同样，尽管对组蛋白乙酰化具有强烈的兴趣，但是核组蛋白乙酰转移酶（HAT）的纯化长期困扰着研究人员。通过一种 HAT 活性的凝胶内分析方法的发展终于有了突破，这种分析方法从四膜虫（*Tetrahymena*）大核中鉴定了第一种核 HAT（Brownell and

Allis 1995；Brownell et al. 1996)。当对编码该 HAT 的基因进行克隆并测序后，发现它与酵母基因 *GCN5* 同源，该基因以前被鉴定为一种转录共激活因子（Georgakopoulos and Thireos 1992；Marcus et al. 1994)。就像 SWI/SNF 的情况，*GCN5* 的鉴定打开了克隆和表征人 *GCN5* 同源物，以及鉴定和表征许多其他以前已发现与转录活化相关联的 HAT 的闸门。自这些早期研究以来，研究人员已经鉴定了组蛋白去乙酰酶、组蛋白激酶、组蛋白甲基化酶，以及组蛋白去甲基化酶等，所有这些酶都对染色质结构和转录调控的复杂动力学发挥作用。已经有了这些酶中许多酶的重组酶，或者已经发表了从不同细胞源纯化它们的方案。下面一节的重点是监测和定量酶活性的生化分析，目的是为转录分析或者表征这个大家族中的新成员做准备。

染色质修饰酶的生化分析

*组蛋白乙酰转移酶。*HAT 活性的检测和表征相对简单，且需要很少的试剂。所有以前鉴定的 HAT 都可以用游离组蛋白作为底物，并且有部分重组 HAT（p300/CBP、SRC1、PCAF）在核小体上有活性。然而，一般来说，最佳的核小体 HAT 活性需要多亚基 HAT 复合物内的其他蛋白质，如内源性酵母 GCN5 HAT 复合物 SAGA（Spt-Ada-GCN5-乙酰转移酶）的 Ada2 和 Ada3 亚基（Balasubramanian et al. 2002)。从鸡红细胞纯化得到的组蛋白是一个很好的内源性组蛋白来源，因为它们含有极少的乙酰化赖氨酸。如果组蛋白纯化自其他来源，建议从所有的缓冲液中删除去乙酰化酶抑制剂（如丁酸钠和曲古抑菌素)，以便内源性去乙酰化酶在纯化期间去除大部分的乙酰化赖氨酸。然而，重组组蛋白是首选，因为它们缺乏可能掩盖酶活性的内源性赖氨酸乙酰化。对于核小体底物，为八聚体转移重建准备的短的寡核小体阵列就很好（见上文)。此外，也可以合成组蛋白氨基端肽作为 HAT 分析的底物。通过合成含有各种取代等位基因的衍生物，可证明合成肽对快速鉴定被作为修饰靶标的赖氨酸残基很有用。

检测和定量 HAT 活性的分析依赖于放射性乙酰基由乙酰辅酶 A（CoA）向组蛋白底物的转移。我们通常使用 [^{3}H] 乙酰辅酶 A（约 5 Ci/mmol)，但也可以使用 [^{14}C] 乙酰 CoA。[^{3}H] 乙酰 CoA 在−20℃可稳定保存至少 1 年，而且较 [^{14}C] 乙酰 CoA 便宜。在我们的分析中，我们使用 3.3 μmol/L [^{3}H] 乙酰 CoA 和 4.7 μmol/L 未标记乙酰 CoA，最终得到 8 μmol/L 的乙酰 CoA，该值接近或超过了被表征的乙酰转移酶的 K_m（Langer et al. 2001；Berndsen et al. 2007)。这为未知酶的分析提供了一个很好的起点，但应滴定 CoA 的浓度，以确定酶活性在线性范围内的浓度。

液相 HAT 分析为检测 HAT 活性提供了最灵敏的方法。反应包含组蛋白底物（100～1000 nmol/L)、乙酰 CoA（8 μmol/L)、不同剂量的酶、10 mmol/L 丁酸钠（抑制污染性组蛋白去乙酰化酶)，以及低盐（小于 100 mmol/L NaCl）缓冲液（pH 7.0～8.0)。通过将等分反应液点样到磷酸纤维素（P-81，Whatman）过滤盘（15mm×15mm）上，来终止反应时间点。迅速用 50 mmol/L 碳酸钠（pH 9.0）溶液冲洗过滤盘，除去未反应的乙酰 CoA 并风干，随后通过闪烁计数量化 [^{3}H] 乙酰 CoA 的掺入。实验中应始终包括无酶对照，以便确定 [^{3}H] 乙酰 CoA 结合的背景水平。如同所有的酶检测，必须小心确保时间点是在反应的线性范围内。我们更喜欢低于 10 min 的时间点，以精确计算在不同组蛋白和酶浓度下的初始速度（Fry et al. 2004)。

液相 HAT 分析在检测总乙酰转移酶活性上表现优异，但它不会产生关于哪个组蛋白多肽已被修饰的信息。这个信息的获得，要通过在 15% SDS-PAGE 凝胶上电泳部分液相 HAT 分析产物，以及通过荧光自显影术检测^{3}H 标记的组蛋白。这种方法不是很敏感，因此，我们通常在这种反应中使用较高浓度的组蛋白（大于 1000 nmol/L），以便几天之内获得适宜的凝胶曝光。

基于液相和凝胶的分析提供了关于 HAT 比活性和底物特异性的信息，但研究人员如何确定哪些残基被修改了呢？一个策略涉及以肽为底物的液相 HAT 分析，该分析中特定的赖氨酸残基被丙氨酸取代。这种分析确定哪些赖氨酸对 HAT 活性是关键的，这些赖氨酸可能等同于也可能不等同于修饰位点。另外，可通过标记蛋白的氨基末端微测序来确定^{3}H 标记乙酰基团掺入的位点（详见 Mizzen et al. 1999）。但是，这种方法依赖于一种可分析放射性蛋白样品的肽测序设施。一个更常见的策略是用未标记的乙酰 CoA 进行 HAT 反应，然后通过质谱分析产物，鉴定标记的肽和修饰位点。

*组蛋白去乙酰化酶。*组蛋白赖氨酸乙酰化是一个动态的标记，该标记长期以来被认为在体内显示迅速的可逆性。1996 年，Schreiber 及其同事纯化并鉴定了第一个组蛋白去乙酰化酶（HDAC），一种酵母转录调节因子类似物——Rpd3（Taunton et al. 1996）。自那时以来，在酵母和哺乳动物中鉴定出 HDAC 大家族，这些酶往往同时显示出组蛋白和赖氨酸位点特异性，并在很多情况下，它们的活性与转录抑制通路相联系。已经开发了若干策略用于体外 HDAC 的活性分析，每种策略各有优点和缺点。有几种方法通过从标记组蛋白肽或者多肽底物中释放［^{3}H］乙酸来监测 HDAC 活性。在所有这些分析中，通过加酸（0.1 mol/L HCl，0.16 mol/L 乙酸）终止反应，用乙酸乙酯提取［^{3}H］乙酸，有机相以闪烁计数量化。数个公司（如 Sigma-Aldrich 和 BIOMOL International）也销售使用荧光或显色底物的去乙酰化酶检测试剂盒。需要注意的是，HDAC 的酵母 SIR2 和 sirtuin 家族与其他 HDAC 不同，因为它们是唯一需要以烟酰胺腺嘌呤二核苷酸（NAD）作为 HDAC 活性辅助因子的 HDAC（Imai et al. 2000；Vaziri et al. 2001）。但在其他方面，SIR2 和 sirtuins 的分析涉及与其他 HDAC 酶相同的策略。

放射性 HDAC 分析在如何准备底物上有所不同。早期的方法涉及培养的 HeLa 细胞或鸡胚红细胞的［^{3}H］乙酸盐代谢标记，随后进行组蛋白的酸抽提纯化。当然，为了保持乙酰基，所有的缓冲液中必须添加组蛋白去乙酰酶抑制剂，如丁酸钠和曲古抑菌素。这种方法的优点是组蛋白乙酰化发生在生理相关位点上，但缺点是标记的组蛋白具有较低的比活性，这会限制 HDAC 活性的检测。这种方法已经成功地被许多研究组，包括 Grunstein 及其同事（Carmen et al. 1996），用于从酵母粗制裂解物中纯化数种 HDAC 酶。

体内标记的替代方法是将单个的组蛋白多肽或合成的组蛋白肽化学乙酰化（Inoue and Fujimoto 1969）。这种方法生成高比活性的标记底物，但标记完整的多肽会生成不处于生理位置上的乙酰赖氨酸。氨基末端组蛋白肽的化学乙酰化是首选，而且，在这种情况下，乙酰化事件对生理修饰会具有代表性。使用肽底物的唯一缺点（这对所有使用组蛋白修饰酶的研究都是适用的）是，在一个折叠的组蛋白或组蛋白-组蛋白对的生理环境下，乙酰化赖氨酸未提交给酶。

第三种［^{3}H］乙酸盐标记策略涉及单个组蛋白或组蛋白的重组 HAT 体外乙酰化。

例如，重组 HAT1 会以高比活性乙酰化组蛋白 H4 氨基末端结构域内的全部 4 个赖氨酸，而重组 GCN5 会乙酰化 H3 氨基末端结构域中的多数赖氨酸。之后通过柱层析或 HPLC，从反应组分中纯化标记的组蛋白（详细的方法，见 Wade et al. 1999）。这种技术没有大的缺点。

使用 [^{3}H] 乙酰-赖氨酸底物的 HDAC 分析需要投入大量的时间和资源。HAT 可能需要纯化，且需要准备并纯化放射性组蛋白，而 HDAC 分析需要^{3}H-标记蛋白的有机抽提。此外，某些 HDAC 显示出了敏感的位点特异性，因此如果希望评估一个特定残基的去乙酰化，该活性可能不容易用多位置标记的底物进行分析。有一种非常简单的监测 HDAC 活性的方法，它依赖于对特定乙酰化组蛋白赖氨酸残基特异的商业化抗体的充足供应。在这种情况下，从细胞来源纯化或从体外 HAT 反应产生非放射性的乙酰化组蛋白。当从细胞纯化组蛋白时，在细胞裂解之前，细胞中加入丁酸钠和曲古抑菌素，促进体内组蛋白超乙酰化，之后在所有后续缓冲液中要包含这些 HDAC 抑制剂。随后将超乙酰化的组蛋白与纯化的或粗制的 HDAC 制剂一起孵育，并将反应产物用 SDS-PAGE 和 Western 分析，以监测特定乙酰化赖氨酸表位的时间依赖性和 HDAC 依赖性损失。

组蛋白甲基化酶及去甲基化酶。核心组蛋白可被在赖氨酸残基（主要是 H3K4、K9、K36、K79 及 H4K20）上单、双或三甲基化或在精氨酸残基（主要是 H3R2、R17、R26 及 H4R3）上单、双甲基化。在 2004 年以前，人们普遍认为，组蛋白赖氨酸和精氨酸残基甲基化是一个不可逆转的标记，这种标记只有通过组蛋白的蛋白质水解或通过细胞分裂从细胞染色质中稀释才能消除。然而，在过去数年中，已经鉴定了两个组蛋白赖氨酸去甲基化酶大家族——LSD1/BHC110 和 JmjC 组蛋白去甲基化酶（Klose and Zhang 2007）。有趣的是，JmjC 家族的其中一个成员，即 JMJD6，也被证明可以催化 H3R2 及 H4R3 去甲基化反应（Chang et al. 2007）。此外，Kouzarides 及其同事们发现，一些甲基化的精氨酸残基也可以通过肽酰基精氨酸脱亚胺酶 4（PADI4）转换成瓜氨酸（Cuthbert et al. 2004）。因此，如组蛋白乙酰化一样，组蛋白甲基化也受到动态调控。

一般用于分析组蛋白甲基化酶和去甲基化酶的策略与那些用于分析 HAT、HDAC 的策略十分相似。对于组蛋白甲基转移酶（HMT）分析，在酶反应中用 *S*-腺苷-L-[甲基-^{3}H] 甲硫氨酸（[^{3}H] SAM，约 80 Ci/mmol）代替 [^{3}H] 乙酰 CoA。[^{3}H] 甲基基团的掺入可以通过 P81 滤纸（液相 HMT 分析）结合或在 15% 的 SDS-PAGE 上电泳反应产物并进行荧光自显影来监测（见图 13.7）。正如 HAT 分析那样，组蛋白底物的选择对 HMT 分析也很关键。对于每一种酶，研究人员应同时测试核心组蛋白和核小体阵列。例如，鉴定甲基化 H4K20 的 Set8 HMT 被推迟了数年，直到发现它只能利用核小体底物（Fang et al. 2002）。此外，一些 HMT，如 PRC2 复合物，偏好甲基化从内源性来源（如小牛胸腺或 HeLa 细胞细胞核）分离的核小体阵列，而不是重组组蛋白组装的阵列（Kuzmichev et al. 2002）。这种增强的活性可能是由其他刺激 HMT 活性的组蛋白标记的存在造成的。

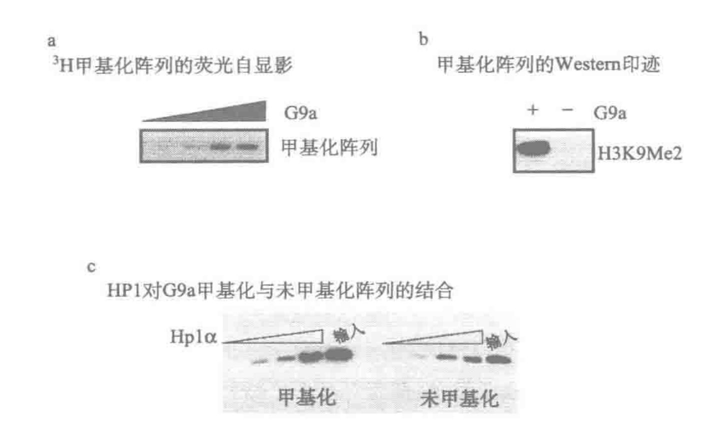


图 13.7　核小体阵列的 G9a 甲基化增强 HP1 的结合。在含有^{3}H 标记 SAM 作为甲基供体的反应中，随着重组 G9a 组蛋白甲基转移酶量的增加，生物素化的 G5E4T-5S 阵列（见图 13.9 说明）被甲基化。(a) 荧光自显影检测组蛋白 H3 的甲基化。(b) 以 H3K9me2 特异性抗体进行 Western 印迹，检测组蛋白 H3 的甲基化。(c) 将甲基化或未甲基化的阵列与重组 HP1α 孵育，再用链亲和素磁珠捕获阵列，最后通过 Western 印迹检测阵列结合的结合型 HP1α 的量。

监测组蛋白去甲基化酶活性的策略也类似于用于 HDAC 的策略。通常，去甲基化酶分析使用从细胞来源（即 HeLa 细胞、小牛胸腺或鸡红细胞）分离的核心组蛋白，去甲基化酶的活性通过使用甲基赖氨酸特异性抗体的 Western 印迹来监测。也可以使用肽底物，通过 MALDITOF/TOF 分析去甲基化反应产物，确定哪些甲基基团已被去除。不像 HDAC，组蛋白去甲基化酶的 JmjC 家族成员需要新的辅助因子。这些酶使用基于羟化的反应机制从赖氨酸上去除甲基基团，并且这种反应需要 Fe（II）（50 μmol/L）和 α-酮戊二酸盐（1mmol/L）（Tsukada et al. 2006；Whetstine et al. 2006）。Fe（II）通常以 $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$ 加入反应。因此，经常将铁螯合剂如 EDTA 和去铁敏（DFO）用作酶抑制剂。显然，用于纯化和分析这些酶的缓冲液应无 EDTA。相比之下，LSD1/BHC110 家族的成员使用黄素依赖性氧化反应来脱去赖氨酸上的甲基，这种反应可以产生甲醛（Shi et al. 2004）。需要注意的是，黄素辅因子，即黄素腺嘌呤二核苷酸（FAD），与酶共纯化，即使从细菌中制备也是如此。甲醛的形成也可以在一种适合动力学分析的光谱分析中测定（详见 Shi et al. 2004）。

*DNA 甲基转移酶。*当组蛋白标记是动态调控时，DNA 中胞嘧啶的甲基化是一个稳定的标记，它可以传递给后代，并且和长期的转录沉默有关。DNA 甲基化是通过 DNA 甲基转移酶（DNMT）进行的，它将甲基从 SAM 转移到 5′位胞嘧啶残基上，这种 5′位胞嘧啶残基通常处于 CpG 二核苷酸中（Klose and Bird 2006）。有两类 DNA 甲基转移酶，包括：①从头转移酶 DNMT3a 和 DNMT3b，它们能修饰未甲基化的 DNA；②维持转移酶 DNMT1，尽管它也可以修饰未甲基化的底物，但它更喜欢甲基化半甲基化的 DNA 底物（Jair et al. 2006）。DNMT1 被认为是复制叉的一种成分，在那儿它可能通过

多重细胞分裂在传播 DNA 甲基化模式中发挥关键作用（Rountree et al. 2000；Esteve et al. 2006）。从头转移酶在早期胚胎发育中起作用，以建立甲基化的结构域。由于这些酶既可以建立又可以传播与转录活性相关的甲基化状态，因此它们代表了那些在发育期间控制稳定表观遗传转录记忆的因子。

DNMT 的体外活性分析相对简单。三种酶都能够从细菌中表达和纯化，并有监测甲基转移酶活性的简单的分析方法。正如 HMT 的分析，DNMT 可以利用 [^{3}H] SAM 作为辅助因子将 [^{3}H] 甲基基团转移到 DNA 上，如 poly（dI-dC）。掺入反应后将反应组分结合到 DE81 纸（Whatman）上，用碳酸氢铵洗涤，再闪烁计数（Pradhan et al. 1999），此法与液相 HAT 和 HMT 分析相似。

作为核小体定位和染色质结构的位点特异性探针，细菌 DNA 甲基转移酶也变得越来越有用。特别是，Kladde 及其同事描述了用于体内外核小体群体和单个分子中核小体定位的甲基转移酶方法（Kilgore et al. 2007；另见第 9 章）。就像限制酶和其他 DNA 结合蛋白，核小体组装抑制 MTase 的甲基转移酶活性。然而，不像限制酶或 DNA 结合蛋白，在大多数序列中，MTase 目标位点的频率很高（通常是 CpG 或 GpC 二核苷酸），这使得核小体定位具有出色的分辨率。最有用的是细菌酶（如 M. SssI 和 M. CviPI），它们分别修饰 CpG 和 GpC 二核苷酸内的 5′位胞嘧啶，因为通过基于 PCR 的亚硫酸测序，可以很容易地检测 DNA 中的 5-甲基胞嘧啶。

ATP 依赖性染色质重塑酶

有四类不同的 ATP 依赖性染色质重塑酶，SWI/SNF、ISWI、INO80 和 Mi-2/CHD，每个组由若干多亚基复合物组成，这些复合物都包含一个催化性 ATP 酶亚基，该亚基属于 DNA 依赖性 ATP 酶 SWI2/SNF2 亚家族（最近的综述，见 Cairns 2005；Durr et al. 2006）。因此，这些酶中的每一个酶都有相当的 ATP 水解能力（每个复合物每分钟水解 200～800 个 ATP），并且显著的 ATP 水解需要核酸辅助因子的存在。而 SWI/SNF 的家族成员在 ATP 酶分析中可以以几乎相同的效率使用双链 DNA、单链 DNA 或核小体为辅因子；ISWI 组的成员需要一种核小体辅因子才能激活显著的 ATP 水解；INO80 和 Mi2/CHD 组的成员可以使用任一类型的辅助因子，但它们似乎更喜欢核小体（Boyer et al. 2000）。一些研究表明，ATP 水解伴随着 DNA 的移位，这会产生超螺旋扭转，并导致核小体重定位成顺式（Cairns 2007，见下文）。因为所有的核小体重塑活性都需要 ATP 水解，所以我们经常使用辅助因子刺激的 ATP 酶活性确定不同重塑酶的比活性或计算不同制剂中活性酶的浓度。

ATP 酶分析。监测 ATP 水解的一个简单方法是，通过在 20 mmol/L 磷酸中将剩余的 ATP 结合在活性炭上，并用闪烁计数定量上清液中的放射性，来测量 [γ-^{32}P] ATP 释放的 [^{32}P] 焦磷酸盐的量（Cote et al. 1994）。不幸的是，这种方法在酶缺失时一般有很高的背景，并且该方法受阻于一个事实，即只能测量反应的一个产物。因此，我们一般使用聚乙烯亚胺（PEI）-纤维素板上的薄层色谱（TLC）来定量 [γ-^{32}P] ATP 和^{32}P 的释放量（Logie and Peterson 1999）。这种方法可以得到重复性很好的结果，而且适合详细的动力学分析。Cairns 等（1994）也提出了一个简单的比色 ATP 酶分析方法，其中无机磷的形成是通过向反应中添加酸性孔雀石绿-钼酸钠溶液进行

定量的。

对于重塑酶的 SWI/SNF 家族成员（如酵母 SWI/SNF 和 RSC 及哺乳动物 BAF 和 PBAF），“染色质重塑”是指染色质底物中大量的体外 ATP 依赖性变化，包括单核小体中组蛋白-DNA 接触的瓦解、顺式和反式组蛋白八聚体的移动、环状微染色体中负超螺旋的丢失、H2A/H2B 二聚体驱逐，以及转录因子和限制酶对核小体 DNA 可及性的增加（综述，见 Smith and Peterson 2005）。相比之下，ISWI、Mi-2/Chd 和 INO80 家族的酶的普遍特点是，它们能够调动顺式核小体或增强限制酶对核小体阵列的可及性。有趣的是，SWR1 酶，一个 INO80 家族成员，未显示调动顺式核小体，但其活性似乎专用于 ATP 依赖性 H2A/H2B 二聚体与 Htz1/H2B 二聚体的交换（Mizuguchi et al. 2003）。

*限制酶可及性分析。*作为 ATP 依赖性核小体重塑活性的初步测量，我们一般使用限制酶可及性分析（Logie and Peterson 1997，1999）。这种分析基于一个事实，即核小体组装呈几个数量级地抑制限制酶切割，而 ATP 依赖性重塑可增强剪切速度 10～30 倍。因此，这通常是一种非常敏感的对重塑活性的测量。原则上，在该类分析中可以使用单核小体、线性核小体阵列或环状核小体阵列。然而，因为单核小体定位可影响某些酶的活性（见下文），核小体阵列底物是首选。

虽然几乎所有的 DNA 片段都可被重建成核小体阵列，并且限制酶切割效率可被监测，但我们发展了一种定量重塑分析，这种分析使用由 11 个 5S rDNA 核小体定位序列串联阵列构成的 DNA 模板（Logie and Peterson 1997，1999）。此模板的主要特点是中心重复被修饰，目的是在定位的核小体中心附近列入唯一的 *Sal*I/*Hinc*II 限制酶识别位点（208-11-Sal 模板）。这种 DNA 片段一般通过盐透析、八聚体转移或快速稀释法重建成核小体定位阵列。当这些重建的核小体阵列暴露于 *Sal*I（或 *Hinc*II）时，切割动力学呈两个阶段。第一阶段非常迅速，其中 25%～50%的阵列被切割，切割速度非常接近裸 DNA。第一阶段切割的这部分阵列代表了群体中中心重复上不含核小体的阵列的比例，这种阵列是由选择性的核小体位置或模板亚饱和造成的。第二阶段慢得多，代表了 *Sal*I/*Hinc*II 位点被定位的核小体阻塞。一般来说，在缺乏重塑酶的条件下对阵列预切割 20 min，以排除第一阶段的切割。当之后将重塑酶和 ATP 添加到反应后，第二阶段的限制酶切割被显著增强。为了开始实验，推荐使用等摩尔比例的核小体重塑酶和核小体（每个阵列 11 个酶分子），以检测较弱的活性。酵母 SWI/SNF 酶反应通常使用每 11 mer 阵列一个 SWI/SNF 的比例，且更低的 SWI/SNF 浓度已证明可以催化重塑反应（Logie and Peterson 1997）。这种二酶耦合反应已被用于测量 SWI/SNF 重塑反应的几个动力学参数（Logie and Peterson 1997），以及转录激活因子招募的 SWI/SNF 活性（Yudkovsky et al. 1999）。这种二酶耦合反应已被用于直接比较 5 种不同酶的重塑活性，这 5 种酶分别为酵母 SWI/SNF、人 SWI/SNF、CHRAC、Mi-2 和 ACF（Boyer et al. 2000）。

*核小体移动分析。*目前已测试的所有的 ATP 依赖性染色质重塑酶（Swr1 可能例外）都利用 ATP 水解产生的能量沿 DNA 顺式移动核小体。7 个不同重塑复合物在单核小体底物上的直接比较表明，每一个复合物都有不同的核小体定位特征（Rippe et

al. 2007)。例如，ACF 和 Mi-2 酶偏爱从 DNA 片段的中心向末端移动核小体，而 ISWI 倾向于从中间向 DNA 末端移动核小体。其他几项研究已经表明，SWI/SNF 和 RSC 重塑酶可以移动核小体离开 DNA 片段的两端，以便小于 147 bp 的 DNA 保持与组蛋白八聚体的结合 (Kassabov et al. 2003)。这些差异归因于不同 ATP 酶亚基及酶特异性非催化亚基的组合。单核小体移动分析基于这样一个事实，即定位于 DNA 片段的中心的核小体在聚丙烯酰胺凝胶中迁移的速度比向 DNA 两端定位的核小体慢。例如，重建在 350 bp Hsp70 启动子片段上的单核小体产生多达 5 个具有不同迁移率的核小体复合物，它代表了选择性的平移定位（见图 13.2）。如果将此单核小体制剂与重塑酶孵育（通常酶与核小体等摩尔比），产物由 PAGE 分析，会发现种类的分布已被改变，或者产生了具有新迁移率的复合物。值得注意的是，对于某些重塑酶（如 SWI/SNF），我们发现，在载样前加入过量的 DNA 对于酶-核小体复合物的解离和取得分离的核小体带型十分必要。这些实验操作起来很简单，它们可用于粗略的动力学分析。此外，可从凝胶中分离出不同种类的单核小体，并进行如前所述的 *Exo*III 核小体位置作图（图 13.2）。

除了顺式移动核小体，重塑酶 SWI/SNF 家族的成员（如酵母 SWI/SNF/和 RSC）也可以置换组蛋白八聚体并将它们反式转移给 DNA 片段（Lorch et al. 1999, 2006）。这种活性在大多数情况下是非常低效的，尽管酵母 RSC 复合物在被测试的酶中似乎是置换能力最强的酶（图 13.8）。此外，多重激活因子蛋白与相关核小体的结合或激活因子对 SWI/SNF 的招募可以增强这一反应（Owen-Hughes et al. 1996; Gutierrez et al. 2007）。八聚体的置换分析在几个方面与核小体的移动分析不同。一般情况下，单核小体被组装在较短的、放射性标记的、接近核心颗粒长度（147～160 bp）的 DNA 片段上，并且应将重建物组装起来，以便保留少量裸 DNA。重塑反应的一个关键加入是过量的未标记受体 DNA。当重塑酶置换组蛋白八聚体时，受体 DNA 将捕获被置换的组蛋白八聚体（图 13.8）。最有效的受体 DNA 包含一个或更多的高亲和性核小体结合位点拷贝，如 601 核小体定位序列。在其他方面，该分析与核小体移动反应类似。在与单核小体底物和受体 DNA 孵育后，反应用 PAGE 分析。可以通过放射性标记游离 DNA 的 ATP 依赖性增加，监测成功的八聚体置换，关键是确保这种反应是 ATP 依赖性的，因为有几个因素可以使体外单核小体稀溶液不稳定。例如，我们发现，对于使用纳摩尔浓度核小体的重塑反应，添加牛血清白蛋白（BSA，100 μg/ml）是该反应的一个重要组分，因为核小体在低蛋白质浓度下可自发解离（Godde and Wolffe 1995）。

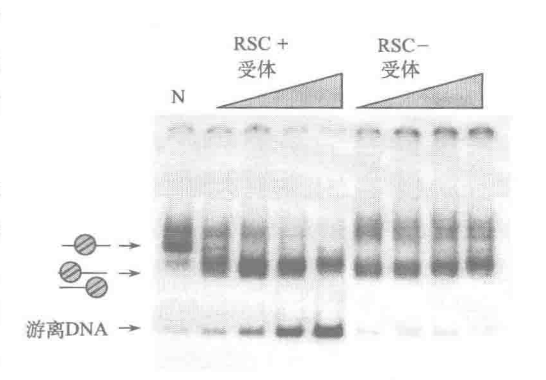


图 13.8 RSC 介导的组蛋白八聚体驱逐需要受体 DNA。RSC 复合物重塑前（泳道 N）和重塑后中央定位单核小体的自然凝胶分析。在 ATP 存在的情况下，RSC 复合物移动了单核小体，并产生了定位于 DNA 片段任一端的产物。当受体 DNA 加入反应时，RSC 还介导组蛋白八聚体的驱逐，导致游离 DNA 的出现。

二聚体驱逐和置换分析。除了顺式与反式移动核小体，有些 ATP 依赖性重塑酶（如 SWI/SNF 和 SWR1）可以取代一个或两个 H2A/H2B 二聚体，或催化经典的 H2A/H2B 二聚体与带有 H2A 变体的二聚体发生交换。例如，在一个需要 Swi3p 亚基组蛋白结合结构域的反应中，在 MMTV 启动子片段上组装单核小体，酵母 SWI/SNF 可以从该单核小体上替换 H2A/H2B（Yang et al. 2007）。此外，在一个需要 Nap1 或 Chz1 组蛋白分子伴侣的反应中，SWR1 复合物可以以 Htz1/H2B 二聚体取代经典的 H2A/H2B 二聚体（Luk et al. 2007）。就人 SWI/SNF 而言，这种二聚体置换反应似乎在体内 MMTV LTR 的转录激活中发挥着关键作用（Yang et al. 2007），并且 SWR1 对 Htz1 在核小体内的沉积必不可少。在酵母细胞中，这些核小体环绕着几乎每个 RNAPII 启动子（Raisner et al. 2005；Zhang et al. 2005）。

已经发展了几种监测 ATP 依赖性 SWI/SNF 复合物置换 H2A/H2B 二聚体的分析方法。Bruno 和 Owen-Hughes 描述了一个利用重组组蛋白 H2A 的简单分析方法，其中重组组蛋白在其羧基末端结构域中设计的半胱氨酸残基上用一个荧光基团［俄勒冈绿（Oregon green）］共价标记（Bruno et al. 2003b）。在该分析中，重组八聚体用俄勒冈绿标记，然后标记的重组八聚体通过盐透析组装到一个未标记的 DNA 片段上，从而产生荧光单核小体。这种标记反应非常简单且显得有效。之后在重塑反应中孵育单核小体，反应产物用 PAGE 分析。为了检测荧光单核小体，我们使用带有 465 nm 激发滤光片和 535 nm 发射滤光片的柯达图像扫描仪（Yang et al. 2007）。通常，测量单核小体复合物中的 ATP 依赖性荧光损失。在某些情况下，加入 H3/H4 四体的阵列可能通过使其担任置换二聚体的受体而促进该反应（Bruno et al. 2003a），尽管我们已经发现当使用完整的酵母 SWI/SNE 时，这种做法是非必需的。俄勒冈绿分析的一个缺点是荧光基团的检测需要在重塑反应中有高浓度的单核小体（30～100 nmol/L 核小体），它将快速耗尽重塑酶库存。俄勒冈绿分析的一个可能容易的替代方法是利用体外染色质免疫沉淀（ChIP）分析监测 H2A 或 H2B 的 ATP 依赖性损失（Yang et al. 2007）。然而，这种分析严重受限于商业化 H2A 和 H2B 抗体的缺乏。

建立一个二聚体交换分析需要几个额外的试剂。第一，通过使用生物素化核苷酸的 Klenow 反应，在一端或两端以生物素标记一个 DNA 片段，然后通过使用重组组蛋白或纯化自细胞的组蛋白八聚体的盐透析，将该标记 DNA 片段组装成核小体阵列。第二种试剂是纯化的 H2A/H2B 或 H2A 变体/H2B 二聚体，这些二聚体将用于分析交换活性。例如，在 SWR1 分析中，从酵母细胞中分离纯化 Htz1/H2B 二聚体（使用含有 Htz1-FLAG 融合蛋白的菌株），或准备为重组组蛋白。当从酵母中纯化时，Htz1/H2B 二聚体与 Nap1 或 Chz1 组蛋白伴侣结合，这种伴侣的活性会极大地刺激交换反应（Luk et al. 2007）。Chz1 也可以制备为重组蛋白加入反应。因此，SWR1 依赖的交换反应包含生物素化的核小体阵列、Htz1/H2B 二聚体、Chz1 或 Nap1 伴侣（通常为具有二聚体的复合物）、ATP 及 SWR1 重塑酶。在不同的时间点，核小体被捕获到链亲和素磁珠上，磁珠用 0.4 mol/L KCl 缓冲液洗涤，除去非特异性结合的组蛋白，并且通过使用商业化 Htz1 抗体的 Western 印迹，量化被交换到阵列上的 Htz1 的量。可以预见，此分析比其他重塑分析需要投入更大的精力，这是因为必须为每套分析滴定 H2A/H2B 二

聚体、组蛋白分子伴侣和重塑酶的用量。此外，为了遵循经典 H2A/H2B 二聚体的交换，倘若缺乏商业化 H2A 抗体，就必须纯化表位标记形式的 H2A（含有重组 H2A-FLAG 载体）。

以染色质模板进行体外转录

正如在本章的前面几节所详细讨论的，体外转录研究通常使用两个常见类型的 DNA 模板。模式核小体阵列模板含有一个二核小体或三核小体长度的启动子区域，该区域被嵌入在一个核小体定位序列（通常是 5S rDNA 基因的串联拷贝）的阵列中。在这种类型的模板中，可以控制核小体位置，从而使关键转录调控因子的结合位点或核心启动子被核小体包围。另外，使用基于提取物的组装系统或重组 ACF/NAP1 系统，可将 DNA 大片段或环状质粒 DNA 组装成均匀间隔的核小体阵列。这些染色质模板不包含定位的核小体，并且因此许多因子的结合位点将是可及的，而且染色质不会阻碍这些初始步骤。

总的来说，设计染色质模板体外转录的研究包含很多与裸 DNA 相同的考虑。必须确定 Pol II 和通用转录因子的来源、量化它们的活性，以及优化缓冲液条件。必须纯化相关的序列特异性激活因子或抑制因子，同时必须将结合位点整合进 DNA 模板。然而，在转录分析中，有几种使用染色质模板的不同方法。一种方法是确定激活因子是否促进染色质组装过程中的转录。例如，激活因子单独或在通用机器存在时，与 DNA 模板一起在胚胎提取物中孵育，这种胚胎提取物使模板组装成染色质模板（或纯化的 ACF/NAP1 重组系统）。如果观察到活化，可能设想两种机制。在其中一种机制中，激活因子可能封阻启动子进入染色质，从而被动地促进基础转录复合物的组装。另外，激活因子可能通过与转录机器的相互作用直接刺激复合物组装。

另一种方法是将染色质预组装到模板上，然后加入激活因子和通用机器。如果激活因子刺激转录，则可以设想两种可能的机制。激活因子可能到达预组装成染色质的启动子，并通过募集染色质重塑酶积极去除染色质，这种重塑可能允许通用机器被动结合。或者，激活因子可能促进染色质重塑，并积极推动转录复合物的组装。已知有些激活因子积极并独立地募集染色质重塑和通用转录机器的组分。例如，GAL4-VP16 直接将 SWI/SNF 复合物和组蛋白乙酰化酶募集到 DNA 上，同时也与转录激活因子、其他共激活因子，以及通用转录因子直接相互作用，以刺激转录。

固定化模板分析/添加分析的顺序

已经描述了很多例子，在这些例子中，一个基因的转录激活涉及调控事件的一个线性顺序。例如，一个基因特异性激活因子可能首先结合到一个或多个位于 RNAPII 依赖性启动子区上游的位点。然后，激活因子可能招募组蛋白乙酰转移酶复合物，该复合物乙酰化核小体附近启动子中的赖氨酸残基。之后，激活因子可能招募 ATP 依赖性重塑酶（如 SWI/SNF），乙酰化的赖氨酸残基可稳固它们与启动子核小体的结合。SWI/SNF 重塑可以移动或者驱逐堵塞转录起始位点的核小体，并且该事件随后将导致起始前复合物（PIC）的组装。当然，这只是一个特定的例子，也有可能是多重可能事件按顺序一起控制转录激活（或抑制）。

有一个强大的策略可使研究者能够仔细分析调控途径并确定不同步骤之间的相互依

存关系，这一策略就是固定化核小体模板分析的使用，加上添加分析的顺序。该策略的关键是在一个 DNA 片段上重建核小体，但该 DNA 片段的一端必须用一个或多个生物素化核苷酸标记，一般通过 Klenow 反应实现这种标记。然后，可将这种生物素化的核小体阵列固定在链亲和素耦合的树脂上。对于生物素化核小体阵列的固定，首选磁珠，因为与各种类型的琼脂糖珠相比，磁珠具有较低的非特异性蛋白结合特性。

一旦固定了核小体模板，就可将单独的调控组分与模板反应，可用加入各种因子的洗涤液去除蛋白质、辅助因子，或交换缓冲液。例如，如果加入激活因子，可通过磁珠上模板的捕获除去过剩的因子，之后用缓冲液温和洗涤。还可以在染色质重塑或组蛋白修饰反应中孵育该模板。一旦完成该反应，则捕获了模板，并通过包含大约 400 mmol/L NaCl 的缓冲液的洗涤清除重塑/修饰酶。另外，可用含有位点特异性组蛋白修饰的重组组蛋白绕过对各种修饰酶的需要。在选择性加入后，可加入 RNAPII 和通用转录因子，以监测转录能力。

建立重塑酶的募集

所有染色质重塑和修饰酶都有对 DNA 和染色质的非特异性亲和性。因此，这种酶与核小体模板的孵育将导致随机的核小体重塑或修饰。然而，在体内，这些事件一般都靶向于选定的位数不多的核小体或一个核小体结构域，并且这种“靶向”活性通常是由序列特异性调控蛋白指导的。尽管重塑酶（如 SWI/SNF）或组蛋白乙酰转移酶（如酵母 SAGA）可以直接和酸性转录激活结构域作用，但这些相互作用是相对低亲和力的。因此，一般无法通过简单地降低酶浓度靶向重塑/修饰活性。唯一的例外是 p300 组蛋白乙酰转移酶，它显示为相当强劲的靶向 HAT 活性的激活因子（图 13.9）。对于许多酶，已经成功的一种方法是向重塑/修饰反应补充不包含相关激活因子结合位点的竞争性核小体（Neely et al. 1999；Yudkovsky et al. 1999）。因此，当一个激活因子预结合到测试核小体阵列时，重塑/修饰酶靶向于这个阵列，而不是靶向于竞争物。如果从反应中消除激活因子，重塑/修饰酶将与所有的核小体同等互作。一般情况下，要滴定竞争物浓度，以免在缺乏预结合激活因子的情况下发生固定化模板的重塑和修饰。

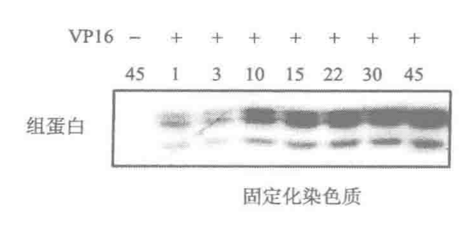


图 13.9 G5E4T-5S 阵列上由 Gal4-VP16 激活因子募集的 p300 介导的组蛋白乙酰化。重组核小体经盐透析重建在 G5E4T-5S DNA 模板上，该模板包含两组 5 个串联拷贝的 208 bp 5S rDNA 核小体定位序列，两组定位序列中间夹着一个二核小体长度的 DNA 片段，该片段含有 5 个 E4 启动子上游的 Gal4-结合位点。DNA 片段的两端在 Klenow 反应中用生物素标记 dUTP 生物素化，同时将生物素化的核小体阵列固定在链亲和素磁珠上。将 p300 HAT 加入到含有或缺乏 Gal4-VP16 激活因子的 [^{3}H] 乙酰 CoA 组蛋白乙酰化反应中。在指定的时间（min）停止反应，用样品缓冲液洗涤并洗脱固定化模板，通过 SDS-PAGE 分离组蛋白。用荧光自显影术检测乙酰化。注意，对于 p300 来说，强烈的乙酰化有赖于激活因子。

体外 ChIP 分析对监测由激活因子靶向的重塑/修饰酶相当有用（Vignali et al. 2000）。对于组蛋白乙酰转移酶来说，也已经使用体外 ChIP 来划定核小体修饰从结合的激活因子处往上游或下游究竟能扩张多远。如典型的 ChIP 分析那样，可以用甲醛诱导蛋白质与蛋白质及蛋白质与 DNA 的交联。除了通过超声剪切 DNA，也可以通过 MNase 消化的限制性酶切将核小体阵列片段化。然后，以相关抗体进行免疫沉淀，回收单个的核小体，再用 PCR 定量。

监测染色质凝聚在转录分析中的作用

如上所述，核小体阵列可根据单价和二价阳离子的作用经历复杂的折叠等级（Fletcher and Hansen 1996）。对于组装在包含头尾相接重复核小体定位序列，如 5S NPE 或合成 601 序列的 DNA 模板上的核小体阵列，已经很好地表征了这些折叠转变。在低盐缓冲液（如 TE）中，这些阵列以伸展纤维的状态存在，这些纤维在电子显微镜下观察看起来像“串珠”结构，而 1～2 mmol/L $MgCl_2$ 的加入会诱导凝缩纤维的形成，这种纤维被认为像生理状态下的 30 nm 纤维。在较高浓度二价阳离子（大于3 mmol/L）中，阵列之间相互作用，形成非常大的寡聚物，看起来可能像直径大于 30 nm 的大的间期染色质纤维。一个典型的核小体阵列的转录反应包含 50～75 mmol/L NaCl 和3～6 mmol/L $MgCl_2$，这种条件可能诱导 30 nm 纤维或更大的寡聚状态的形成。最近，Tremethick 及其同事通过在分析超速离心机中的沉降速度，监测了伴随着核小体阵列凝聚的体外转录（Zhou et al. 2007）。这些作者发现，在转录缓冲液中核小体阵列形成大的寡聚物，此外，还发现影响染色质凝聚程度的组蛋白改变会对体外转录造成严重后果。

似乎很清楚，研究人员必须思考策略，以评估染色质凝聚在目的基因的调控属性中的作用。理想的情况下，一个策略是以改变的组蛋白重建核小体阵列，这些组蛋白在高级结构的形成中有缺陷。例如，缺乏 H4 氨基末端尾巴结构域或在 H4K16 中包含乙酰基的组蛋白无法在体外形成 30 nm 样纤维，而且不太容易形成寡聚物（Shogren-Knaak et al. 2006）。然而，H4 氨基端这样的改变也可能扰乱其他调控事件。另外，组蛋白 H2A 中特定酸性残基的突变也会破坏体外的阵列折叠（Zhou et al. 2007）。最后，众所周知，30 nm 样纤维的形成需要一个连续的核小体阵列。因此，组装轻微亚饱和的阵列（即有一个或多个空白的阵列）也会妨碍高级结构的形成（Hansen and Lohr 1993）。手中有了这样的工具，研究人员就可以评估一个调控因子是否在破坏染色质高级折叠中发挥作用或者在其他步骤中起作用。

技术

方案 13.1 鸡红细胞组蛋白八聚体制备

可以从各种来源的细胞中纯化核心组蛋白，包括果蝇胚胎、组织培养的 HeLa 细胞、小牛胸腺或鸡红细胞。我们经常用鸡红细胞作为细胞组蛋白来源，因为获得大量纯化的鸡血非常容易，并且纯化的组蛋白具有低水平的翻译后修饰。也可以从同一个细胞

样品中纯化连接组蛋白 H1 和 H5。这里所述的 H1 和 H5 的纯化方案可以作为一个可选步骤。请注意，禽组蛋白的氨基酸序列与人类组蛋白相同。200 ml 血液的核心组蛋白产量通常为 50 mg。在这个程序的结尾也提供了制备 H3/H4 四聚体和 H2A/H2B 二聚体的方案，可作为可选步骤。已经在 TE/2mol/L NaCl 缓冲液中 4℃成功保存组蛋白原液超过一年，在−20℃已经保存了数年（方案由美国科罗拉多州立大学 Jeff Hansen 惠赠）。

时线与安排

从鸡红细胞纯化核心组蛋白共需三天时间。血液必须立即处理。我们发现，向血液添加肝素可增强血液在短期内的稳定性。第二次 MNase 消化（步骤 7）之后和羟基磷灰石色谱步骤（步骤 8）之前，样品可在冰上储存几天。接收血液前，应该准备好所有缓冲液并冷却。

提纲

步骤 1：准备缓冲液并冷却（2 h）。
步骤 2：制备细胞核（1 h）。
步骤 3：微球菌核酸酶消化 I（2 h）。
步骤 4：CM Sephadex C-25 分级分离（3.5 h）。
步骤 5：过夜透析。
步骤 6：连接组蛋白 H1/H5 准备（可选）。
步骤 7：微球菌核酸酶消化 II（2～5 h）。
步骤 8：羟基磷灰石色谱分析（6 h）。
步骤 9：H2A/H2B 二聚体和 H3/H4 四聚体的制备（可选）。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

琼脂糖凝胶（1%）
缓冲液 1
- 1×储存缓冲液
- 2.5 mmol/L EDTA
- 0.5 mmol/L EGTA

缓冲液 2
- 1×储存缓冲液
- 2.5 mmol/L EDTA
- 0.5 mmol/L EGTA
- 0.5% NP-40

缓冲液 3

1×储存缓冲液

缓冲液 4

15 mmol/L 柠檬酸钠（柠檬酸）〈!〉

150 mmol/L NaCl

缓冲液 5

0.1 mol/L 磷酸钾（pH 6.7）

0.1 mol/L NaCl

缓冲液 6

0.1 mol/L 磷酸钾（pH 6.7）

1.1 mol/L NaCl

缓冲液 7

0.1 mol/L 磷酸钾（pH 6.7）

2.2 mol/L NaCl

$CaCl_2$（0.1 mol/L）〈!〉

鸡血（新鲜，250 ml；Pel-Freeze，33132-1）

考马斯蓝染料〈!〉

ddH_2O

EDTA（0.25 mmol/L 和 25 mmol/L）

洗脱缓冲液

0.1 mol/L 磷酸钾（pH 6.7）

2.2 mol/L NaCl

乙醇〈!〉

溴化乙锭〈!〉

5×HTP 上样缓冲液

0.5 mol/L 磷酸钾缓冲液

0.5 mol/L NaCl

微球菌核酸酶（MNase；Worthington）

Milli-Q H_2O

NaCl（0.35：1.6 mol/L 溶于 10 mmol/L Tris〈!〉，pH 8.8）

NaCl（0.6：1.2 mol/L 溶于 10 mmol/L Tris〈!〉，pH 8.8）

NaCl（5 mol/L）

NaOH（0.1 mol/L）〈!〉

pBR322 *Msp*I 分子质量标准物

苯甲基磺酰氟（PMSF）〈!〉储液（0.1 mol/L）

蛋白酶抑制剂（抑酶肽〈!〉，亮肽素〈!〉，PMSF〈!〉）

SDS 凝胶（18%分离胶，6%堆积胶）〈!〉

5×SDS 上样染料〈!〉

250 mmol/L Tris-HCl（pH 6.8）

10% SDS〈!〉
5% β-巯基乙醇〈!〉
0.02%溴酚蓝〈!〉

起始柱缓冲液（10 mmol/L Tris〈!〉，pH 8.8，0.35 mol/L NaCl，或 10 mmol/L Tris〈!〉，pH 8.8，0.6 mol/L NaCl）

2×储存缓冲液
30 mmol/L Tris 碱〈!〉
30 mmol/L NaCl
120 mmol/L KCl〈!〉
1 mmol/L 亚精胺〈!〉
0.3 mmol/L 精胺〈!〉
680 mmol/L 蔗糖

TE 缓冲液
10 mmol/L Tris（pH 7.8）〈!〉
0.25 mmol/L EDTA
0.1 mmol/L PMSF〈!〉

Tris（1 mol/L，pH7.8）〈!〉

洗涤缓冲液
0.1 mol/L 磷酸钾（pH 6.7）
0.1 mol/L NaCl

设备

Amicon 浓缩器，带 XM50 膜
烧瓶（250 ml）
离心机（GSA 和 HB6 转头）
CM-Sephadex C-25 柱色谱仪
透析设备
电泳仪
Eppendorf 管（1 ml）
吸液管（25 ml）
分光光度计
37℃水浴
水纯化系统（Millipore）

程序

步骤 1：准备缓冲液并冷却

所有缓冲液 4℃储存。重要：所有缓冲液中的 PMSF（储存 0.1 mol/L）都须在立

即使用前加入，终浓度为 0.1 mmol/L（羟基磷灰石色谱使用的缓冲液见步骤 7）。

2×储存缓冲液（250 ml 鸡红细胞染色质需 4 L）

	分子质量/(g/mol)	所需量/g
30 mmol/L Tris 碱	121.14	14.54
30 mmol/L NaCl	58.44	7.01
120 mmol/L KCl〈!〉	74.56	35.79
1 mmol/L 亚精胺〈!〉	145.25	0.628 ml 0.925 g/ml
0.3 mmol/L 精胺〈!〉	348.20	0.418
680 mmol/L 蔗糖	342.30	931.06

用盐酸调 pH 为 7.5；用 Milli-Q 水定容到 4 L。

缓冲液 1（250 ml 鸡红细胞染色质需 2 L）

1×储存缓冲液	1.0 L	2×原液
2.5 mmol/L EDTA	10 ml	0.5 mol/L EDTA
0.5 mmol/L EGTA	5.0 ml	0.2 mol/L EGTA

用 Milli-Q 水定容到 2 L。

缓冲液 2（250 ml 鸡红细胞染色质需 4 L）

1×储存缓冲液	2.0 L	2×原液
2.5 mmol/L EDTA	20ml	0.5 mol/L EDTA
0.5 mmol/L EGTA	10ml	0.2 mol/L EGTA
0.5% NP-40	20 ml	100% NP-40

用 Milli-Q 水定容到 4 L。

缓冲液 3（250 ml 鸡红细胞染色质需 2 L）

1×储存缓冲液	1.0 L	2×原液

用 Milli-Q 水定容到 2 L。

缓冲液 4（250 ml 鸡红细胞染色质需 2 L）

	分子质量/(g/mol)	所需量/g
15 mmol/L 柠檬酸钠（柠檬酸）〈!〉	294.1	8.82
150 mmol/L NaCl	58.44	17.53

用盐酸调 pH 为 7.2，用 Milli-Q 水定容到 2 L。

步骤 2：制备细胞核

这个步骤包括鸡红细胞的清洗和裂解，以获得干净的细胞核颗粒。每种缓冲液在使用前立即加入 PMSF。在所有步骤中始终将溶液保持在冰上。

1. 提前准备 0.1mol/L PMSF 储液：溶解 0.436 g PMSF 在 25 ml 乙醇中，在铝箔包裹瓶子，－20℃储存。

2. 将 250 ml 新鲜鸡血分装加入 4 个 250 ml 的瓶子中，每瓶 62.5 ml。

3. 加 0.1 mol/L PMSF 到缓冲液 4（1 ml/L；终浓度为 0.1 mmol/L）中。用此缓冲液加满 250 ml 鸡血瓶，用 GSA 转头 4℃，3000 r/min（1464g）离心 250 ml 瓶

子 5min。

注意：沉淀块非常松软，小心清干上清。

4. 用缓冲液 4 重悬沉淀，GSA 转头 4℃，3000 r/min（1464 *g*）离心 250 ml 瓶子 5 min，重复三次。在此阶段，努力从瓶中只清干黄色液体。

5. 缓冲液 1 加 0.1 mol/L PMSF（1 ml/L；终浓度为 0.1 mmol/L）。沉淀重悬在缓冲液 1 中，并用 GSA 转头 4℃，3000 r/min（1464 *g*）离心 5 min，重复两次。

6. 缓冲液 2 加 0.1 mol/L PMSF（1 ml/L；终浓度为 0.1 mmol/L）。沉淀重悬在缓冲液 2 中，并用 GSA 转头 4℃，3000 r/min（1464 *g*）离心 5 min，重复 3～4 次。

在这步结尾，应出现很白的细胞核沉淀。应从重复重悬中消除所有来自血红素的红色或粉红色颜色。重悬沉淀时，用 5 ml 吸管反复吹打细胞核沉淀，该方法有助于破碎细胞核团块和消除红色或粉红色颜色。

7. 缓冲液 3 加 0.1 mol/L PMSF（1 ml/L；终浓度为 0.1 mmol/L）。沉淀重悬在缓冲液 3 中，并用 GSA 转头 4℃，3000 r/min（1464 *g*）离心 5 min。在第一次离心后，检查沉淀是否呈深色斑块，并彻底重悬大块沉淀使之消失。重复重悬和离心两次。

最后的洗涤非常关键，因为它从混合物中去除 EDTA。如果一些 EDTA 仍然被保留，MNase 消化将无法工作。

8. 细胞核重悬于 60 ml 缓冲液 3（全部合并入一个瓶子）中。避免较大的体积，因为下一步需要一定的浓度。

9. 定量 DNA：细胞核稀释在 0.1 mol/L NaOH 中（5 μl 细胞核加 995 μl 0.1 mol/L NaOH），测定 A_{260} 值。(A_{260}×200）/20 ＝ mg/ml；mg/ml×60 ml ＝总 mg。

步骤 3：微球菌核酸酶消化 I

这部分准备工作涉及有限的 MNase 消化、细胞核裂解和核膜的消除。

1. 用双蒸水水化 MNase，使原液为 45 000 U/ml（此原液 500 倍稀释应为消化提供适当的酶浓度，90 U/ml）。

2. 用缓冲液 3 和 0.1 mol/L $CaCl_2$ 原液溶液稀释重悬的细胞核，使 DNA 终浓度为 3 mg/ml，而 $CaCl_2$ 的终浓度为 1 mmol/L（0.1 mol/L 原液的 100 倍稀释）。

3. 细胞核放置在 37℃水浴，涡漩混匀 5 min 加热。

细胞培养瓶很适合漩涡混匀。

4. 加 MNase 原液，使 MNase 浓度为 90 U/ml（30 U/mg DNA）。37℃水浴再涡漩混匀 5 min。

5. 置于冰上，加入 0.25 mol/L EDTA 至终浓度为 2.5 mmol/L（1 ml/100 ml）。

6. 悬液在 GSA 转头中 4℃，6500 r/min（约 6800 *g*）离心 5 min。

7. 重悬沉淀在 1 体积 0.25 mmol/L EDTA、0.1 mmol/L PMSF 中，相当于 1/4 消化体积（0.25 mol/L EDTA 和 0.1 mol/L PMSF 1000 倍稀释在水中）。

避免更大的体积，因为在 CM-Sephadex 被加入步骤中需要一定的浓度。

8. 破碎沉淀，在冷藏室中搅拌 1 h。

此步裂解细胞核。

9. 沉淀混合液在 GSA 转头中 4℃，6500 r/min（约 6800 *g*）离心 20 min，收集上清。

10. 定量 DNA：用水稀释上清，读取 A_{260}值（5 μl 上清液加 995 μl ddH_2O）。(A_{260} × 200)/20 = mg/ml；mg/ml×ml 上清液＝总 mg DNA。

步骤 4：CM-Sephadex C-25

这部分准备工作是为了从寡核小体中去除组蛋白 H1 和 H5。

1. 来自步骤 3 的上清液用水稀释到 5 mg/ml，但稀释液中包括 1 mol/L Tris（pH 7.8）、25 mmol/L EDTA 和 5 mol/L NaCl，稀释后其终浓度为 10 mmol/L Tris、0.25 mmol/L EDTA 和 0.35 mol/L NaCl。

   执行这一步是为了让染色质回到缓冲液中。在第 3 步中裂解细胞核的 MNase 消化是在 H_2O 和 EDTA 中，没有缓冲液。用 TE 替代 1 mol/L Tris 会导致 EDTA 的终浓度为 0.5 mmol/L，而不是 0.25 mmol/L，但这不要紧，它将有助于螯合所有的 Ca^{2+}。这一点很重要，因为有些 MNase 依然存在在制备物中。
2. 加 0.1 mol/L PMSF，维持制备物中 0.1 mmol/L PMSF 浓度。
3. 加 CM-Sephadex，每毫升染色质使用 60 mg 干树脂（即 12 mg 树脂/ml 染色质）。
4. 在冷室中搅拌 CM-Sephadex 混合物 3 h，这段时间为下一步浸泡透析管。

   如果要保存 CM-Sephadex 用于回收 H1/H5（见下面可选步骤 6），此时不应搅拌。而要在 3 h 孵育期间振荡混合 CM-Sephadex 浆液。
5. 混合物在 HB6 转头中 4℃，6500 r/min 离心 30 min。用吸管小心取出上清液。

   如果 CM-Sephadex 将用于回收 H1 和 H5，此时，在上清液透析期间，必须接着进行回收步骤，参考步骤 6（H1/H5 制备）。

步骤 5：过夜透析

1. 用 2 L TE 缓冲液 4℃过夜透析来自步骤 4：5 的上清液。
2. 当天上午，用新鲜 TE 透析上清液（约 2 h）。

步骤 6：连接组蛋白 H1/H5 准备（可选）

这步的第一部分与通过 CM-Sephadex 步骤制备组蛋白八聚体相同。如果同时制备 H1 和 H5 与八聚体，应该从 CM-Sephadex 上洗脱 H1 和 H5，而将寡核小体上清在 TE 中透析过夜。参考 Garcia-Ramirez 等（1990）的原始论文。

这步将组蛋白 H1/H5 作为两个独立的峰从 CM-Sephadex 洗脱下来。以下问题很重要。

- pH 是至关重要的。
- 初步洗涤是至关重要的（除去散布在 Sephadex 中的二聚体和寡核苷酸）。
- 梯度的坡度很重要。降低盐梯度（0.6～0.2 mol/L NaCl）的坡度也可协助分离 H1 峰前污染性二聚体。如果使用 0.6～1.2 mol/L NaCl 梯度，在完成这些盐梯度洗脱之后，考虑用 1.5 倍体积的 1.6 mol/L NaCl 洗柱，以洗脱任何可能仍然结合在柱上的

物质。

• 所有缓冲液添加 PMSF，使其至终浓度为 0.1 mol/L，并尽快操作以避免蛋白质降解。

1. 在起始柱缓冲液（10 mmol/L Tris，pH 8.8，0.35 mol/L NaCl 或 10 mmol/L Tris，pH 8.8，0.6 mol/L NaCl）中洗涤来自步骤 4：5 的 CM-Sephadex 颗粒 2～3 次，每次洗涤后，用 HB6 转头 4℃，6500 r/min 离心 10 min（同样继续对 CM-Sephadex 温柔操作）。

2. 在 CM-Sephadex 柱（2.5 cm×约 26 cm）顶上铺一层洗涤的 CM-Sephadex。CM-Sephadex 柱事先已经过选择的起始柱缓冲液的平衡。

3. 用 1L 盐梯度溶液（0.35～1.6 mol/L NaCl，10 mmol/L Tris，pH 8.8 或 0.6～1.2 mol/L NaCl，10 mmol/L Tris，pH 8.8）洗脱。收集 120 滴（4 ml）馏分。为整个梯度都收集馏分，通过读取馏分的 A_{230} 找到峰。对于 0.35 mol/L NaCl 开始的梯度而言，预计 H1 在馏分 90 左右洗脱，而 H5 在馏分 140 左右洗脱。

4. 各馏分加蛋白酶抑制剂。

抑酶肽　10 μg/ml 终浓度

亮肽素　10 μg/ml 终浓度

PMSF　0.1 mmol/L 终浓度

从 1000×原液（1 μl/ml）加入，以免稀释馏分。

5. SDS-PAGE 电泳，分离胶 18%，浓缩胶 6%。确定组蛋白质量。从步骤 6：3 收集并在步骤 6：4 处理的各馏分中分别上样 1 μg 组蛋白。考马斯亮蓝染色。

6. 汇集来自步骤 6：3 的包含连接组蛋白的馏分，分装并储存在 4℃。

步骤 7：微球菌核酸酶消化 II

这一步将寡核小体消化成小碎片，包括三聚体、二聚体和单体。如果将大的寡核小体载样到羟基磷灰石（HTP）柱上，则八聚体的洗脱峰会太分散。

1. 染色质制剂的浓度应仍然大约为 5 mg/ml，通过两次读取 A_{260} 值仔细检查。

2. 加 0.1 mol/L $CaCl_2$ 至终浓度 1 mmol/L。

3. 如下准备测试消化：

染色质（5 mg/ml）	30 μl（150 μg）
MNase（4500 U/ml；原液的 10 倍稀释）	1 μl（4.5 U 或 0.03 U/mg 染色质）

4. 37℃孵育混合物，并在 5 min、10 min、20 min 和 30 min 时间点采样。在每个时间点，取 1 μl（约 5 μg），并加入到一个含有 99 μl 2.5 mmol/L EDTA 的 Eppendorf 管中。

5. 所有时间点采样完成后，为 1%琼脂糖凝胶准备样品，如下所示。

a. 每个时间点取 10 μl（约 0.5 μg）。

b. 加 2 μl 5×载样染料/5%SDS。

c. 37℃加热 30 min，1%琼脂糖凝胶 80 V 电泳 1 h；用 pBR322 *Msp*I 作标准。

d. 溴化乙锭染色并观察。

6. 以消化效果最佳的时间点建立一个大规模消化反应。通过加入 0.25 mol/L EDTA至终浓度 2.5 mmol/L 终止反应。

7. 消化反应在 Amicon 浓缩器（使用 XM50 膜）中浓缩；浓缩至约 1/10 体积（即约 50 mg/ml）。

即使在 MNase 消化后，浓度也应大约为 5 mg/ml。期望浓度大约为 50 mg/ml，这是因为在 HTP 柱载样之前，一份消化产物与 5×HTP 柱缓冲液混合。

步骤 8：羟基磷灰石色谱

这一步去除 DNA，以产生游离八聚体。

1. 按 Bio-Rad 的说明小册子灌注 HTP 柱。

2. 制备 5×HTP 载样缓冲液（0.5 mol/L 磷酸钾缓冲液，0.5 mol/L NaCl，具有所需的工作浓度：0.1 mol/L 磷酸钾，pH 6.7，0.1 mol/L NaCl）。检查磷酸钾试剂瓶，弄清分子质量。这将视水合状态而定。

溶质	g（无水）	g（三水合）
K_2HPO_4（二元）	17.42	22.82
KH_2PO_4（一元）	13.62	
NaCl	5.84	

在 Milli-Q 水中溶解，定容到 0.2 L（200 ml）。

3. 制备使 500 ml 1×洗涤缓冲液（0.1 mol/L 磷酸钾，pH 6.7，0.1 mol/L NaCl）。取 100 ml 载样缓冲液，并加 Milli-Q 水到 0.5 L（500 ml）。

4. 制备洗脱柱缓冲液（0.1 mol/L 磷酸钾，pH 6.7，2.2 mol/L NaCl）。

K_2HPO_4（二元）	8.70 g（11.42 g $KH_2PO_4 \cdot 3H_2O$）
KH_2PO_4（一元）	6.80 g
NaCl	64.28 g

在 Milli-Q 水中溶解，定容到 0.5 L（500 ml）。

5. 将 0.8～4 ml 的浓缩核小体与 0.2～1 ml 的 5×HTP 载样缓冲液混合。从 HTP 柱顶去除洗涤缓冲液，并小心将样品铺到柱顶。当样品运行进入柱后，轻轻地铺一层1×洗涤缓冲液，并连接储液槽。用 0.5～1 体积的 0.1 mol/L 磷酸钾（pH 6.7）和0.1 mol/L NaCl 缓冲液（洗涤缓冲液）洗涤。然后，连接含 1×洗脱缓冲液（0.1 mol/L磷酸钾，pH 6.7，2.2 mol/L NaCl）的储液槽，并收集 180 滴馏分（每个馏分约6 ml；约 30 滴/ml）。柱的流速应大约为 1 ml/min。收集 80 个组分，尽管峰应在介于组分 35 及 45 的某处洗脱。

6. 通过读取每个馏分的 A_{230} 值定位峰。用一份 50 μl 小等份读取吸光度，然后将它丢弃。1 mg/ml 的八聚体具有 4.3 的 A_{230} 值。因此，每个馏分的 A_{230} 除以 4.3 将给出该馏分单位为 mg/ml 的浓度。

7. 各馏分加蛋白酶抑制剂。

抑酶肽　　10 μg/ml 终浓度

亮肽素　　10 μg/ml 终浓度

PMSF　　0.1 mmol/L 终浓度（从 1000×原液加，以免稀释馏分）

8. SDS-PAGE 电泳，分离胶 18%，浓缩胶 6%。确定组蛋白质量。各组分分别上样 4 μg 组蛋白，100 V 跑胶 1～1.5 h。考马斯亮蓝染色。

9. 组蛋白 4℃储存，各馏分分装成 6 个 Eppendorf 管（每管 1 ml）。

10. 为了重复利用 HTP 柱，请从顶部去除几厘米 HTP 树脂，用新鲜的 HTP 取代。用几个柱体积的 0.5 mol/L 磷酸钾缓冲液（pH 6.7）洗涤柱，以洗脱 DNA。如果柱子将长时间不用，应将其储存在含 0.02% NaN_3（0.77 g/L）的 1×洗涤缓冲液中。临用前，用无 NaN_3 的 1×洗涤缓冲液平衡。

步骤 9：H2A-H2B 二聚体和 H3/H4 四聚体的制备（可选）

本制备方案的第一部分与步骤 6 制备组蛋白八聚体相同。但是，HTP 层析却不同。在八聚体制备中，柱缓冲液含有 2.2 mol/L 的 NaCl 浓度，它会破坏所有带正电荷的组蛋白和带负电荷的 DNA 之间的离子相互作用。所有组成八聚体的组蛋白将被洗脱，而 DNA 仍然结合。在此程序中，以含有渐增 NaCl 浓度的缓冲液进行分步洗脱。H2A/H2B 二聚体在较低的 NaCl 浓度中洗脱，H3/H4 四聚体在较高 NaCl 浓度中洗脱。

1. 准备如下缓冲液：

缓冲液 5（0.1 mol/L 磷酸钾，pH 6.7，0.1 mol/L NaCl）

K_2HPO_4（二元）	17.42 g（22.82 g $K_2HPO_4 \cdot 3H_2O$）
KH_2PO_4（一元）	13.61 g
NaCl	5.85 g

加水定容到 1 L。

缓冲液 6（0.1 mol/L 磷酸钾，pH 6.7，1.1 mol/L NaCl）

K_2HPO_4（二元）	17.42 g（22.82 g $K_2HPO_4 \cdot 3H_2O$）
KH_2PO_4（一元）	13.61 g
NaCl	64.3 g

加水定容到 1 L。

缓冲液 7（0.1 mol/L 磷酸钾，pH 6.7，2.2 mol/L NaCl）

K_2HPO_4（二元）	17.42 g（22.82 g $K_2HPO_4 \cdot 3H_2O$）
KH_2PO_4（一元）	13.61 g
NaCl	128.6 g

加水定容到 1 L。

2. 按 Bio-Rad 的说明小册子灌注 HTP 柱。使用一个 1.5 cm×30 cm 玻璃柱，并倒柱约 25 cm 高。在缓冲液 5 中平衡柱。

3. 混合 800 μl 浓缩核小体（约 40 mg）和 200 μl 5×HTP 载样缓冲液（0.5 mol/L 钾磷酸，pH 6.7，0.5 mol/L NaCl）。

4. 让样品流入柱，然后连接缓冲液 5 储液槽。共收集 15 个馏分，每个馏分 90 滴

（约 3 ml）。

5. 切换到缓冲液 6 储液槽，并收集至少 30 个馏分，每馏分各收集 90 滴（约3 ml）。切换到缓冲液 7 储液槽，并收集至少 30 个馏分，每馏分各收集 90 滴（约 3 ml）。二聚体应该洗脱在缓冲液 6 开始后约第 10～第 12 个馏分，四聚体应洗脱在缓冲液 7 开始后第 10～第 12 个馏分。通过读取 A_{260} 值找到峰。

6. 各馏分加蛋白酶抑制剂。

抑酶肽	10 μg/ml 终浓度
亮肽素	10 μg/ml 终浓度
PMSF	0.1 mmol/L 终浓度（从 1000×PMSF 原液加，以免冲淡馏分）

7. SDS-PAGE 电泳，18%分离胶，6%浓缩胶。确定组蛋白质量。每馏分上样约 2 μg组蛋白。

8. 分装馏分，在 4℃保存。

9. 为了重复利用 HTP 柱，从顶部去除几厘米 HTP，代之以新的 HTP。用数个柱体积的 0.5 mol/L 磷酸钾（pH 6.7）洗柱。如果柱子长时间不使用，可将其存放在 1×洗涤缓冲液（如果长期储存，应添加 0.02% NaN_3）中。

附加说明

1. 其他细胞来源。来自 Dignam 和 Roeder 核提取物制备方法（见第 12 章）的细胞核可作为步骤 2：7 中鸡红细胞细胞核的替代品。

2. 八聚体浓度。一般情况下，组蛋白八聚体将以约 2 mg/ml 的浓度从 HTP 柱上洗脱。如果最终的八聚体汇聚液的浓度远远低于 1 mg/ml，则应浓缩组蛋白。八聚体稀溶液似乎不那么稳定，并且重建染色质的效率可能较低。我们通常使用低分子质量截留（4500）的 Centricon 型微量浓缩器。

3. 我们经常在 4℃存放组蛋白八聚体超过 1 年。八聚体也可以透析成包含 50%甘油的 HTP 洗脱缓冲液，并在−80℃保存。

疑难解答

第一次 MNase 消化之前细胞核很难重悬

可能原因：细胞核可能在洗涤步骤已裂解。

解决方案：每个重悬步骤中的操作一定要温柔。不要涡旋，但可用大口径玻璃吸液管轻轻吸打。检查缓冲液配制和玻璃器皿，以确保缓冲液配制正确和玻璃器皿不含残留洗涤剂。

组蛋白从 HTP 柱以宽峰洗脱

可能原因：不彻底的 MNase 消化导致 HTP 柱上上样了很长的寡核小体。

解决方案：可合并和浓缩组蛋白。用琼脂糖凝胶检查一份 HTP 载样，以确认寡核小体大小。

方案 13.2 核小体的盐梯度透析重建

在盐梯度透析方案中，纯化的核心组蛋白与 DNA 模板在含有高浓度 NaCl（2 mol/L）的缓冲液中孵育。由于盐被缓慢透析掉，核小体在 DNA 上自发组装，并且它们沿着 DNA 的平移定位受 DNA 序列的指引。在缺乏 NPE（如 5S rDNA 基因）的情况下，核小体可采取非生理性的、核小体间几乎无间隔的紧密排列结构。可以通过蔗糖梯度分级分离除去剩余的游离组蛋白，以及带紧密排列核小体的模板。在这些反应中组装的染色质常常要经受微球菌核酸酶（MNase）酶切分析（方案由美国康奈尔大学 Lee Kraus 惠赠）。

时线与安排

盐透析重建很容易建立，但很费时间，因为透析缓冲液变化多端。这种重建方法需要花费 1.5～2 天才能完成。所有的透析缓冲液在重建开始前应准备妥当并被冷却。

提纲

步骤 1：准备缓冲液并冷却（1 h）。
步骤 2：建立染色质组装反应（30 min）。
步骤 3：盐梯度透析和样品收集（30～36 h）。
步骤 4：MNase 消化分析（4 h）。
- MNase 消化。
- DNA 样品制备。
- 琼脂糖凝胶电泳。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

琼脂糖凝胶
乙酸铵（2.5 mol/L）
溴酚蓝染料〈!〉
缓冲液 R
　10 mmol/L HEPES-KOH（KOH〈!〉，pH 7.5），其中包含：
　10 mmol/L KCl〈!〉
　1.5 mmol/L $MgCl_2$〈!〉
　0.5 mmol/L EGTA
　10%（*V/V*）甘油
　分装并保存于−20℃。可冻融多次。

透析缓冲液原液

NaCl（5 mol/L）

Tris-Cl（1 mol/L，pH 8.0）

EDTA（0.5 mol/L）

DTT（0.25 mol/L）〈!〉

DNA 模板

乙醇〈!〉

溴化乙锭〈!〉

组蛋白八聚体原液

微球菌核酸酶原液（200 U/ml）

200 U 微球菌核酸酶（Sigma-Aldrich，N5386）溶于 1 ml 含 2.5 μmol/L $CaCl_2$ 的 5 mmol/L 磷酸钠缓冲液（pH 7.0）。

分装后保存在−20℃，可以冻融多次。

MNase 终止液

27.5 μl TE〈!〉

17.5 μl 0.5 mol/L EDTA

5 μl RNase

NaCl 原液（5 mol/L）

酚：氯仿：异戊醇（25：24：1 *V/V/V*）〈!〉

PMSF（0.1 mol/L）〈!〉

PK 消化液

20 mmol/L EDTA

Na^+，pH 8，包含：

0.2 mol/L NaCl

1%（*m/V*）SDS〈!〉

0.25 mg/ml 糖原（Sigma-Aldrich，G0885）。

室温储存。

蛋白酶 K 原液

2.5 mg/ml 蛋白酶 K（USB20818）溶于 TE 缓冲液〈!〉

−20℃保存。可以冻融多次。

无 DNase 的 RNase，以 0.5 mg/ml 蛋白质溶液形式供货（Boehringer Mannheim）

1×Tris/硼酸/EDTA（简称 TBE）〈!〉

TE 缓冲液（1.0 mol/L、0.8 mol/L、0.6 mol/L、0.0025 mol/L）

5×TG 载样缓冲液

50%（*V/V*）甘油

5 mmol/L EDTA（pH 8.0）

0.1%（*m/V*）溴酚蓝〈!〉

设备

Beckman SW41 转头
离心机
透析管（3000 M. W. 截留）
　　或 MWCO（分子质量截留）迷你透析杯（Pierce Biotechnology）
　　或 Slidelyzer 透析卡（Pierce Biotechnology）
Eppendorf 管（1.5 ml）
微量离心机
离心管（硅化，1.5 ml）
SDS-PAGE 电泳仪
分光光度计
SpeedVac 或其他旋转浓缩仪
紫外透射仪

程序

步骤 1：准备缓冲液并冷却

所有缓冲液储存在 4℃。

透析缓冲液（终体积都为 1 L）

试剂原液	TE 1.0 mol/L	TE 0.8 mol/L	TE 0.6 mol/L	TE 0.0025 mol/L
5 mol/L NaCl	200 ml	160 ml	120 ml	0.5 ml
1 mol/L Tris-Cl，pH 8.0	10 ml	10 ml	10 ml	10 ml
0.5 mol/L EDTA	0.5 ml	0.5 ml	0.5 ml	0.5 ml

临用前加入 2 ml 0.1 mol/L PMSF 和 2 ml 0.25 mol/L DTT。

步骤 2：建立染色质组装反应

盐透析重建一般在 100 μg/ml DNA 浓度组装，尽管我们已经使用 50 μg/ml 和 200 μg/ml的浓度成功组装了染色质。重建中使用的核心组蛋白八聚体的用量取决于所期望的核小体占据的最终水平。例如，要组建一个被核小体完全占据的染色质底物，使用的比例为每 200 bp DNA 模板一个组蛋白八聚体。在质量基础上，这相当于每克 DNA 0.793 g 组蛋白八聚体。当 DNA 模板专门由核小体定位序列组成时（如 208 bp 的 5S rDNA 基因串联重复序列），使用比例则为每个 NPE 一个组蛋白八聚体。例如，208 bp 5S rDNA 阵列，使用的质量比为每克阵列模板 0.77 g 八聚体。

请注意，在这个例子中，反应将有 500 μl 的终体积，并包含 50 μg 2.4 kb 的质粒 DNA（1 584 000 Da）。要用核小体饱和该 DNA，每个质粒 DNA 分子需加入 12 个组蛋白八聚体（每 200 bp 一个八聚体；$r=1$），将需要 39.65 μg 八聚体。不过，我们建议

总是准备三种不同的组装反应，分别包含摩尔比（*r* 值）为 0.9、1.0 和 1.1 的组蛋白八聚体和 200 bp DNA。尽管这些反应只有 10%的变化，但要获得一个具有适当组装水平的核小体阵列，这些变化往往是需要的。

步骤 3：盐梯度透析和样品收集

1. 解冻 DNA 模板，复检浓度。
2. 微量离心机离心组蛋白八聚体原液 10 min，除去聚集物。通过读取 230 nm 处的吸光度复检浓度。
3. 在一 1.5 ml Eppendorf 管中混合所需量的组蛋白与 DNA，并调整到 500 μl。轻弹管底混合，不要涡旋。用 5 mol/L NaCl 原液调整 NaCl 浓度到 2 mol/L。
4. 剪下一小块 3000 MWCO 的透析管，用水冲洗。在透析管中建立染色质组装反应。

   另外，对于小型（小于 1 ml）重建，我们已经使用了 10 MWCO 迷你透析杯（Pierce Biotechnology）；或者对于大型重建，使用 Slidelyzer 盒（Pierce Biotechnology）。
5. 透析按逐步方式进行，如下所述。

步骤	透析缓冲液	时间
1	1 L TE + 1 mol/L NaCl	4 h 到过夜
2	1 L TE + 0.8 mol/L NaCl	4 h 到过夜
3	1 L TE + 0.6 mol/L NaCl	4 h 到过夜
4	1 L TE + 0.0025 mol/L NaCl	4 h 到过夜
5	1 L TE + 0.0025 mol/L NaCl	4 h

6. 小心收集样品到 1.5 ml 硅化离心管（组蛋白在 4℃很黏!!!），4℃储存。

   盐透析染色质可存放在 4℃下至少 1 个月。

要除去游离组蛋白，染色质样品可以在大约 12 ml 15%～40%蔗糖梯度上分级分离（每梯度可将约 200 μg 的 DNA 组装成染色质；Beckman SW41 转头 4℃下 30 000 r/min 离心 16 h）。取 300 μl 馏分分析 DNA（琼脂糖凝胶 EB 染色）和核心组蛋白（SDS-PAGE）。含 DNA 和组蛋白（即染色质）的馏分的峰相对而言应无游离组蛋白。

步骤 4：体外组装染色质的微球菌核酸酶分析

MNase 消化分析基于一种事实，即 MNase 优先剪切核小体之间的连接 DNA，而且当酶在极限消化条件下使用时，可由此产生含有不同个数核小体的 DNA 片段。脱蛋白后的 DNA 片段可用琼脂糖凝胶电泳分开，而且，如果已经发生染色质组装，可以看到一个由逐渐增大的 DNA 片段形成的重复模式。此法可用于确定核小体的间距变化，就像连接组蛋白未掺入时看到的那样。

MNase 消化

1. 建立如下染色质样品

a. 在 150 μl 缓冲液 R 中稀释 2 μl 盐透析的染色质样品。

b. 加 4.5 μl 0.1 mol/L $CaCl_2$（这是 MNase 活性所需）到 150 μl 稀释液中，立即轻轻弹试管混合。

c. 准备两个 50 μl 等份（每个含约 300 ng DNA）。

样品编号	样品	MNase 稀释
1	盐透析染色质	1∶20 000
2	盐透析染色质	1∶10 000

重要：在开始消化之前准备好 MNase 终止液!!!（见 B 部分）

2. 加入 5 μl 稀释的微球菌核酸酶（见上表和下表）。对于每个染色质样本，在缓冲液 R 中准备如下微球菌核酸酶稀释，立即使用时准备。

盐透析染色质		
MNase 稀释	MNase 体积	缓冲液 R
1∶400	2 μl 原液	798 μl
1∶10 000	2 μl 1∶400	48 μl
1∶20 000	2 μl 1∶400	98 μl

3. 立即轻弹试管混合。连续启动多个样本，15 s 间隔。

   在该方案中，描述了使用几个不同浓度的 MNase 取得不同程度的消化。然而，恒定浓度的 MNase 也可以使用，但消化的时间可能各不相同。

4. 室温下消化 10 min。

5. 加入 5 μl MNase 终止液，立即轻弹试管混合。连续终止多个样品，15 s 间隔。

   MNase 终止液中的 EDTA 螯合钙离子并抑制反应。

DNA 样本的制备

6. 加入 100 μl PK 消化液，彻底涡旋，37℃孵育 30 min。

   蛋白酶 K 在临用前加入蛋白酶 K 消化液，如下所述。在加入样品之前，将蛋白酶 K 原液与上面列出的溶液按 1∶10（*V/V*）的比值合并（即将蛋白酶 K 原液 1∶10 稀释在列出的缓冲液中）。每次使用都制备新鲜稀释。

7. 用 150 μl 苯酚∶氯仿∶异戊醇抽提。在微量离心机中室温下离心 5 min。

8. 转移水相到新管（避免界面!），并加 15 μl 2.5 mol/L 乙酸铵及 340 μl 乙醇。涡旋，在微量离心机中室温下离心 15 min。

   不要用乙酸钠，它会产生大的油性沉淀。

9. 去除上清液，在旋转浓缩仪（如 SpeedVac）中干燥沉淀。

10. 将沉淀重悬在 10 μl 水中，室温静置 10 min。加 2.5 μl 5×TG 缓冲液并混匀。

    此步后样本可储存在−20℃。

琼脂糖凝胶电泳

11. 在 1×TBE 中用 1%（*m/V*）的琼脂糖凝胶 150 V 电泳样品、一个对照 MNase 消化 DNA 样品，以及 123 bp DNA 梯状条带。溴酚蓝染料迁移到凝胶 2/3 长度时，停止电泳。

12. 凝胶在含 0.75 μg/ml 溴化乙锭的水中室温染色 30 min。紫外透射仪观察 DNA。

凝胶可在水中室温脱色达 2 h，以减少背景荧光。如果需要更好的分辨率，染色后的凝胶需要再稍微电泳一下。

疑难解答

重建后检测不到 DNA

可能原因：核酸酶污染 DNA、组蛋白或缓冲液原液。

解决方案：每个有 DNA 的组分先 37℃孵育 30 min，之后琼脂糖凝胶检测 DNA 的完整性，以此测试核酸酶。也可以将透析缓冲液中的 EDTA 浓度提高到 1 mmol/L。

可能原因：重建中 DNA 聚集，导致沉淀。

解决方案：重新分析 DNA 和组蛋白原液的浓度，以确认以适当的比例组装。因为游离的核苷酸可影响 DNA 的吸光度读数，用琼脂糖凝胶电泳一份 DNA 模板，以确认吸光度读数。同样，用 SDS-PAGE 电泳组蛋白原液与 BSA 浓度标准物，以检查组蛋白原液。

可能原因：用于分析的 MNase 原液浓度太高或终止缓冲液不正确。

解决方案：在没有 MNase 消化的情况下，检查重建物的完整性。对于阵列模板（约 2.5 kb），重建的阵列迁移速率略快于游离 DNA 片段。

核小体重建，但阵列亚饱和

可能原因：DNA 或八聚体组蛋白浓度不准确。

解决方案：通过吸收法和琼脂糖凝胶或 SDS-PAGE 电泳分析重新检查浓度。

可能原因：八聚体原液不具备完全组装的能力。

解决方案：在某些情况下，重组八聚体原液中含有超过 10%的没有组装能力的八聚体，可能由不完全的折叠造成。当使用窝藏各种氨基酸替换的八聚体时，这个比例可能会增加。在这种情况下，简单地增加 r 值（如每 200 bp 重复用 1.3～1.8 个八聚体），以取得更大的重建效率。

方案 13.3　利用重组的果蝇 ACF 和 NAP1 重建核小体阵列

染色质组装程序的目标是从克隆的 DNA 模板和纯化的核心及连接组蛋白出发，以准备扩展的核小体阵列。组装的染色质应在蛋白质含量方面高度明确，并且在周期性和核小体定位方面与从活细胞核分离出来的散装染色质相似。在这里，我们描述如何以 ATP 依赖性方式从环状质粒 DNA 模板和纯化的核心组蛋白组装微型染色体。该系统也可用于从线性 DNA（质粒和 λ）组装微型染色体，也可以整合核心组蛋白之外的蛋白质（连接组蛋白 H1、HMG17 和 DNA 结合转录因子）。染色质组装反应的产物已直接（或纯化后）用于各种分析，来研究转录、DNA 复制、重组和修复。该系统采用纯化的重组果蝇染色质组装因子 ACF 和 NAP1。这些组装因子的纯化方案也提供如下（该方案由美国加利福尼亚大学圣迭哥分校 Jim Kadonaga 和 Dimitri Fyodorov 惠赠）。

时线与安排

将核心组蛋白与 NAP1 在冰上孵育，形成组蛋白-伴侣复合物。然后，通过加入 ACF、ATP 和质粒 DNA（超螺旋、松弛或线性），在 27℃启动染色质组装反应。染色质组装的程度可通过分析环状 DNA 超螺旋进行监测，而染色质的“质量”可通过微球菌核酸酶消化分析进行检测。组装反应可在 1.5～2.5 h 内完成，分析需要约 2 h。

概要

步骤 1：准备缓冲液（1 h）并为染色质组装反应组装主溶液（30 min）。

步骤 2：染色质组装（1.5～2.5 h）和 MNase 消化分析（2 h）。

附录 1：重组果蝇 ACF 的纯化。

附录 2：重组果蝇 NAP1 的纯化。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

ACF 稀释缓冲液：含 0.4 mg/ml 重组人胰岛素（Roche）的洗涤缓冲液 F（配方见下）

琼脂糖凝胶，1×TBE

AM 主混合液

3 μl	0.5 mol/L ATP
30 μl	0.5 mol/L 肌酸磷酸
16.5 μl	双蒸水
25 μl	100 mmol/L $MgCl_2$〈!〉

使用前加入 0.5 μl 5 mg/ml 肌酸激酶。

乙酸铵（2.5 mol/L）

ATP（0.5 mol/L）

牛血清白蛋白（BSA）溶液（2 mg/ml）

缓冲液 R

10 mmol/L K^+-HEPES（pH 7.6）

10 mmol/L KCl〈!〉

1.5 mmol/L $MgCl_2$〈!〉

0.5 mmol/L EGTA

10%（V/V）甘油

4℃储存溶液 24 h。

使用前加入：

10 mmol/L β-甘油磷酸

1 mmol/L DTT〈!〉
0.2 mmol/L PMSF〈!〉
$CaCl_2$（10 mmol/L）〈!〉
经核心组蛋白储存缓冲液透析的核心组蛋白（纯化）（0.3～2.0 mg/ml）
核心组蛋白储存缓冲液
10 mmol/L K^+-HEPES（pH 7.6）
1 mmol/L EDTA
10 mmol/L KCl〈!〉
10%（*V/V*）甘油
4℃储存溶液到 24 h。
使用前加入 1 mmol/L DTT。
肌酸激酶溶液
5 mg/ml 肌酸激酶（Sigma-Aldrich）
10 mmol/L 磷酸钾（pH 7.0）
50 mmol/L NaCl
50%（*V/V*）甘油
分装成 5～10 μl 小等份，－80℃可储存达 2 年。
肌酸磷酸（0.5 mol/L）
106 mg/ml 肌酸磷酸（磷酸肌酸）
20 mmol/L K^+-HEPES（pH 7.6）
调 pH 到 7.0。分装成 0.1～1 ml 小等份，－20℃储存达两年。
松弛 DNA 模板
7.64 μl 质粒 DNA（0.42 mg/ml）
2 μl 10×拓扑异构酶 I 缓冲液
2.36 μl 重组拓扑异构酶 I 工作液
8 μl 双蒸水
EDTA（0.5 mol/L）
乙醇（100%）〈!〉
溴化乙锭〈!〉
糖原终止缓冲液
20 mmol/L EDTA
0.2 mol/L NaCl
1%（*m/V*）SDS〈!〉
0.25 mg/ml 糖原
室温可储存两年。
HEG 缓冲液
25 mmol/L K^+-HEPES（pH 7.6）
0.1 mmol/L EDTA

10% (V/V) 甘油

分装成 0.1～1 ml 小等份，－20℃可储存达 2 年。

KCl (300 mmol/L) 〈!〉

$MgCl_2$ (100 mmol/L) 〈!〉

微球菌核酸酶原液 (200 U/ml)

1.56 mg/ml (200 U/ml) 微球菌核酸酶 (Sigma-Aldrich)

5 mmol/L 磷酸钠 (pH 7.0)

2.5 μmol/L $CaCl_2$ 〈!〉

分装成 0.1～1 ml 小等份，－20℃储存达 1 年。

酚：氯仿：异戊醇〈!〉，(50：49：1 V/V/V)，10 mmol/L Tris-Cl 〈!〉(pH 8.0) 平衡

质粒 DNA，在 TE 缓冲液中 (0.3～2.0 mg/ml)

蛋白酶溶液

1.1 ml 糖原终止缓冲液

55 μl 2.5 mg/ml 蛋白酶 K

PvOH/PEG 溶液

HEG 缓冲液（配方见上）含有：

5% 聚乙烯醇 (M.W. 10 000；Sigma-Aldrich P8136)

5% 聚乙二醇 (M.W. 8 000；Sigma-Aldrich P2139)

分装成 0.1～1 ml 小等份，－20℃储存达两年。

重组 ACF (0.002～0.2 mg/ml)（关于制备，见下面附录 1）

重组 NAP1 (0.5～4.0 mg/ml)（关于制备，见下面附录 2）

RNase A (10 mg/ml) 〈!〉

终止液

55 μl 0.5 mol/L EDTA

11 μl 10 mg/ml RNase A 〈!〉

拓扑异构酶 I（果蝇，重组）

10×拓扑异构酶 I 缓冲液

0.5 mol/L Tris-Cl (pH 7.5)

100 mmol/L $MgCl_2$ 〈!〉

1 mmol/L EDTA

0.5 mg/ml BSA

5 mmol/L DTT 〈!〉

0.1～1 ml 小等份，－20℃可储存到两年。

洗涤缓冲液 F

20 mmol/L Tris-Cl (pH 7.9)

150 mmol/L NaCl

15% (V/V) 甘油

2 mmol/L $MgCl_2$〈!〉
0.2 mmol/L EDTA
0.01%（V/V）NP-40（Nonidet P-40）
4℃储存溶液 24 h。
使用前加入：
1 mmol/L DTT〈!〉
10 mmol/L β-甘油磷酸
0.2 mmol/L PMSF〈!〉
0.5 mmol/L 苯甲脒-HCl〈!〉
2 μg/ml 亮肽素〈!〉
1 μg/ml 抑酶肽〈!〉
二甲苯青染料〈!〉

设备

电泳仪
试管（硅化聚丙烯管，1.5 ml）

程序

步骤 1：为染色质组装准备主溶液

1. 解冻所有缓冲液和蛋白质。
建议将所有的缓冲液平衡至室温，蛋白质必须快速解冻（室温水浴或转移到冰上）并且使用后快速冷冻（液态氮）。ACF、NAP1、核心组蛋白、核酸和微球菌核酸酶可以承受多次冻融循环。

2. 准备 NAP1（见附录 2）和核心组蛋白（NH）的主混合液，在硅化 1.5 ml 离心管中合并以下试剂：

172 μl	HEG 缓冲液
70 μl	300 mmol/L KCl
84 μl	PvOH/PEG 溶液
4.2 μl	2 mg/ml BSA 溶液
6.4 μl	2 mg/ml NAP1
3.03 μl	0.7 mg/ml 的核心组蛋白。

温和涡旋 2～3s。
整个方案使用硅化的 1.5 ml 试管。所提供的配方按 6 个标准反应计算，应该用于进行 5 个或更少的反应（允许不准确的移液）。NAP1 与核心组蛋白的质量比为 5∶1，应足够消除反应中未结合的组蛋白。

3. 向 5 个硅化的 1.5 ml 的离心管中各加入一份步骤 1：2（室温）制备的 56.6 μl NH 混合液小等份。冰上孵育≥20 min，以使组蛋白-NAP1 结合。

4. 准备 ATP 和 Mg^{2+} 的 AM 主混合液。

5. 准备 DNA 模板。30℃松弛 DNA 10 min，室温保存，直到使用。

反应中的拓扑异构酶应 5～10 倍过量于在 30℃，10 min 内彻底松弛超螺旋质粒所需的酶量。纯化后的系统在缺乏拓扑异构酶的情况下也将有效地在超螺旋 DNA 上组装染色质。在后一种情况下，模板必须含有 95%以上的超螺旋 DNA。

6. 在 ACF 稀释缓冲液中准备 ACF 稀释液（2～10 U/μl）。保持在冰上（1 U ACF 等于 22 fmol 蛋白质）。

步骤 2：组装染色质并用微球菌核酸酶分析法分析

1. 开始如下组装反应

a. 向 56.6 μl NH（步骤 1：3）加 1 μl ACF 的稀释缓冲液（2～10 U）

b. 从冰上转移开，平衡至室温。

c. 加 10.5 μl 主混合液（步骤 1：4）和 2 μl 的 DNA 模板（用过量的拓扑异构酶 I 松弛过的 DNA，或超螺旋 DNA；步骤 1：5），立即温和涡旋 2～3s。

d. 让组装在 27℃进行 1.5～2.5 h。

2. 立即使用前，在缓冲液 R 中准备两个微球菌核酸酶稀释液（1∶500 和 1∶1500）。

3. 每个反应加入 17.5 μl 10mmol/L $CaCl_2$。将每个反应分为两等份（“a”和“b”）。以受控方式（在一定的时间间隔，如 15 s），向每个“a”管中加入 5 μl 步骤 2：2 准备的 1∶1500 稀释液，向每个“b”管加入 5 μl 步骤 2：2 准备的 1∶500 稀释液。让每管在室温进行 10 min 消化。

该组装反应也可以用 DNA 超螺旋分析方法监控。通过加入 3 μl 0.5mol/L EDTA 终止 1/4 的 70 μl 的组装反应（17.5 μl 中约 0.177 μg DNA）。如步骤 2：5 进行 DNA 的脱蛋白并沉淀。在 0.8%琼脂糖凝胶上电泳超螺旋和松弛 DNA 样本，并同步跑 1 kb DNA 梯状条带，电泳缓冲液为 1×TBE，电泳至二甲苯青染料带到达凝胶底端 1/3 处。溴化乙锭染色和脱色。

4. 准备终止液（ST）。每管加入 6 μl ST 终止微球菌消化，然后涡旋。让样品在室温下静置 5 min，以消化污染的 RNA。

5. 准备蛋白酶溶液（PR）。每管加 105 μl PR，然后涡旋。37℃消化组蛋白和可溶性蛋白大于 30 min。用 200 μl 酚∶氯仿∶异戊醇（50∶49∶1）抽提样品。用 25 μl 2.5 mol/L 乙酸铵和 475 μl 100%乙醇沉淀 DNA。

6. 进行琼脂糖凝胶电泳。

附录 1：重组果蝇 ACF 的纯化

果蝇 ACF（利用 ATP 的染色质组装和重塑因子）是通过杆状病毒共表达羧基末端 FLAG 标记 Acf1 亚基与未标记 ISWI 亚基准备的。复合物通过 FLAG 免疫亲和层析一步纯化，此程序一般产生 Acf1 和 ISWI 以化学计量结合的复合物。

提纲

步骤 1：Sf9 细胞感染和收获。

步骤 2：准备和分析 ACF 蛋白。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

Acf1-FLAG 和 ISWI 杆状病毒原液，高滴度

目前这些原液唯一的来源是 J. Kadonaga 实验室，University of California，San Diego。

牛血清白蛋白（BSA）标准物（2 mg/ml；Pierce，23209）

稀释缓冲液 F

20 mmol/L Tris-Cl，pH 7.9

10%（V/V）甘油

0.02%（V/V）NP-40（Nonidet P-40）

4℃储存达 24 h。

洗脱缓冲液 F

洗涤缓冲液 F（配方见下），其中包含：

0.4 mg/ml FLAG 肽（Sigma-Aldrich）

0.4 mg/ml 重组人胰岛素（Roche）

立即使用。从 10 mg/ml FLAG 肽原液加 FLAG 肽，然后从 50 mg/ml 胰岛素原液加胰岛素。FLAG 肽原液由 FLAG 肽溶于 STE 缓冲液（100 mmol/L NaCl，10 mmol/L Tris〈!〉，pH 8.0，1 mmol/L EDTA）配成。胰岛素原液由胰岛素溶于 TE 缓冲液〈!〉（10 mmol/L Tris，pH 7.4，1 mmol/L EDTA）配成。

FLAG-M2 树脂（1∶1 V/V 浆液；Sigma-Aldrich）〈!〉，在裂解缓冲液 F 中已平衡

液氮〈!〉

裂解缓冲液 F

20 mmol/L Tris-Cl（pH 7.9）

500 mmol/L NaCl

20%（V/V）甘油

4 mmol/L $MgCl_2$〈!〉

0.4 mmol/L EDTA

溶液在 4℃储存达 24 h。

添加后立即使用：

2 mmol/L DTT（二硫苏糖醇）〈!〉

20 mmol/L β-甘油磷酸盐
0.4 mmol/L PMSF〈!〉
1 mmol/L 盐酸苯甲脒
4 μg/ml 亮肽素〈!〉
2 μg/ml 抑酶肽〈!〉
磷酸盐缓冲液（PBS；冰冷）
Sf9 细胞（对数晚期）悬浮培养（>2Å，约 10^6 细胞/ml）
洗涤缓冲 F
20 mmol/L Tris-Cl（pH 7.9）
150 mmol/L NaCl
15%（V/V）甘油
2 mmol/L $MgCl_2$〈!〉
0.2 mmol/L EDTA
0.01%（V / V）NP-40
4℃存储溶液到 24 h。
添加后立即使用：
1 mmol/L DTT〈!〉
10 mmol/L β-甘油磷酸
0.2 mmol/L PMSF〈!〉
0.5 mmol/L 盐酸苯甲脒
2 μg/ml 亮肽素〈!〉
1 μg/ml 抑酶肽〈!〉

设备

临床离心机，带吊桶式转头
锥形离心瓶（250 ml），适宜临床离心机，或锥形管（50 ml）
锥形离心管（14 ml 和 50 ml；一次性）
培养板（150 mm）
塑料管（15 ml，带盖）
硅化 1.5 ml 聚丙烯管
Sorvall 超速离心机，带 SS-34 转头（或同等物）
飞旋瓶（150 ml 或 500 ml）
Wheaton dounce 匀浆器（“A”杵，15 ml）

程序

步骤 1：感染和收获 Sf9 细胞

1. 感染前几天扩增杆状病毒原液。病毒以细胞培养上清的方式提供，并要在 4℃保

存。扩增病毒，将细胞以 2×10^7 个细胞/板、总体积为 25 ml 适当的含血清培养基的量接种到 150 mm 培养板中，并以 0.1～0.5 的感染复数用病毒感染细胞（10～20 μl 病毒上清/平板）。为了传递一个存储病毒的原液，让感染进行 60 h，收集培养液上清。为了获得高滴度的原液，让细胞持续感染，直到出现明显的细胞裂解（His-NAP1 病毒感染约 72 h，ISWI 和 Acf1-FLAG 病毒感染约 84 h）。从培养板中吸出培养液，储存在灭菌的 50 ml 试管中，4℃黑暗中保存（最长达 12 个月）。

2. Sf9 细胞以 0.5×10^6 个细胞/ml 的密度接种在 150 ml 或 500 ml 飞旋瓶中，生长 2～3 天。以每板 2.5×10^7～3×10^7 个细胞、总体积 25 ml 适当昆虫培养液的量，将 Sf9 细胞接种 5～25 个培养板。让细胞在组织培养罩中定居 20 min，并以 5～10 的感染复数（MOI）用 Acf1-FLAG 和 ISWI 杆状病毒感染细胞。

3. 在感染后 44～46 h，吸出培养液，每板用 10 ml 冰冷 PBS 将细胞从板上洗脱。在 250 ml 锥形瓶或 50 ml 试管中用临床离心机 2000 r/min 4℃离心 5 min。

细胞沉淀可在液氮中冷冻，并储存在－80℃，在进一步处理前可保存数周。

4. 重悬细胞沉淀于 8 ml 的裂解缓冲液 F 中，用 Wheaton dounce 匀浆器瓦解沉淀。使用“A”杵在 30 min 内于冰上进行三个系列的 10 次击打。

步骤 2：准备和分析 ACF 蛋白

1. 在 14 ml 锥形管中 4℃ 下 14 500 *g*（SS-34 转头 11 000 r/min）离心 10 min，沉淀不溶性颗粒物质。将上清与 250 μl FLAG-M2 树脂（在裂解液 F 中平衡的 1∶1 浆液）和 7 ml 稀释缓冲液合并。在 15 ml 带盖聚丙烯管中于摇床上 4℃混合 4 h。

2. 洗涤树脂 4 次，每次使用 12 ml 洗涤缓冲液 F；每次洗涤后先用临床离心机在 4℃下 2000 r/min 离心 3 min，之后吸出上清，再加入洗涤缓冲液 F，颠倒离心管，重悬树脂。

3. 如下述步骤洗脱蛋白质。

a. 向步骤 2：2 沉淀的树脂加入 100 μl 洗脱缓冲液 F，重悬沉淀。

b. 树脂转移到 1.5 ml 硅化离心管中。

c. 冰上孵育 10 min。

d. 最大速度离心 30 s。

e. 将上清转移到另一管中（将与随后的洗脱合并）。

f. 继续在相同的硅化离心管中，重复步骤 2∶3a～e 三次，收集所有洗脱液。

4. 液氮中冷冻分装的蛋白质小等份（20～50 μl），并储存在－80℃。

重组 ACF 稳定达数年，并能承受多次（5～10）冻融循环。每个 150 mm 培养板的 ACF 的一般产量少于 5～10 μg。

附录 2：重组果蝇 NAP1 的纯化

Sf9 细胞感染 HIS6-NAP-1 表达杆状病毒。用 Ni^{2+}-NTA 亲和层析纯化伴侣蛋白，之后在 Source 15Q 树脂（Pharmacia）上进行阴离子交换色谱。

提纲

步骤 1：感染和收获 Sf9 细胞。

步骤 2：通过镍亲和层析纯化 NAP1。

步骤 3：通过阴离子交换层析纯化 NAP1。

材料

注意：以〈!〉标记的材料的恰当处理见附录。

试剂

牛血清白蛋白（BSA）的标准物（2 mg/ml，Pierce，23209）

缓冲液 R

- 10 mmol/L K^+-HEPES（pH 7.6）
- 10 mmol/L KCl〈!〉
- 1.5 mmol/L $MgCl_2$〈!〉
- 0.5 mmol/L EGTA
- 10%（*V/V*）甘油
  - 溶液在 4℃中存储长达 24 h。
- 立即使用前添加：
  - 10 mmol/L β-甘油磷酸
  - 1 mmol/L DTT（二硫苏糖醇）〈!〉
  - 0.2 mmol/L PMSF〈!〉

洗脱缓冲液 H：洗涤缓冲液 H（配方见下），其中含 480 mmol/L 咪唑〈!〉

乙醇（20%，*V/V*）〈!〉

HEGD 缓冲液—0.1 mol/L NaCl

- 25 mmol/L K^+-HEPES（pH 7.6）
- 1 mmol/L EDTA
- 10%（*V/V*）甘油
- 0.1 mol/L NaCl
- 0.01%（*V/V*）NP-40
  - 具有上述组分的溶液 4℃存储到 24 h。
- 添加后立即使用：
  - 1 mmol/L DTT〈!〉
  - 0.2 mmol/L PMSF〈!〉
  - 10 mmol/L β-甘油磷酸

His-NAP1 杆状病毒原液（高滴度）

目前这些原液唯一的来源是圣地亚哥加利福尼亚大学 Kadonaga 实验室。

液氮〈!〉

裂解缓冲液 H

50 mmol/L 磷酸钠（pH 7.6）

0.5 mol/L NaCl

15%（V/V）甘油

20 mmol/L 咪唑〈!〉

0.01%（V/V）NP-40

具有上述组分的溶液 4℃存储到 24 h。

添加后立即使用：

10 mmol/L β-甘油磷酸

0.2 mmol/L PMSF〈!〉

0.5 mmol/L 苯甲脒

NAP1 纯化缓冲液

缓冲液 R（配方见上），其中包含 0.0、0.1 mol/L 或 1.0 mol/L NaCl（添加自 5 mol/L NaCl 原液或固体 NaCl）、0.01% NP-40（添加自 10% V/V 原液）。4℃储存达 24 h。

Ni-NTA 琼脂糖树脂（QIAGEN）〈!〉

磷酸盐缓冲液（PBS；冰冷）

SDS-PAGE 凝胶（8%和 15%）〈!〉

Sf9 细胞（对数晚期），悬浮培养（大于 2×10^6 个细胞/ml）

Source 15Q 树脂（Amersham Pharmacia Biotech）〈!〉

洗涤缓冲液 H

50 mmol/L 磷酸钠（pH 7.6）

100 mmol/L NaCl

20 mmol/L 咪唑〈!〉

15% 甘油（V/V）

0.0 1% NP-40（V/V）

4℃储存溶液达 24 h。

立即使用前加入：

10 mmol/L β-甘油磷酸（添加后立即使用）

0.2 mmol/L PMSF（添加后立即使用）〈!〉

0.5 mmol/L 苯甲脒（添加后立即使用）

设备

离心瓶（250 ml 锥形）适宜用于临床离心机

临床离心机，带吊桶式转头

透析管（12 000~15 000 MWCO）

快速蛋白质液相色谱（FPLC）装置

HR-5 或 HR-b FPLC 柱（Amersham Pharmacia Biotech）

旋转混合仪（颠倒式）
SDS-PAGE 电泳仪
Sorvall 超速离心机，带 SS34 转头（或同等物）
飞旋瓶（500 ml 或 1000 ml）
试管（锥形，15 ml 和 50 ml）
试管（硅化聚丙烯，1.5 ml）
Wheaton Dounce 匀浆器（40 ml 带“A”杵）

程序

步骤 1：感染和收获 Sf9 细胞

1. 在 500 ml 或 1000 ml 转瓶中培养 Sf9 细胞，使其密度大于 2.0×10^6 个细胞/ml 培养液。用培养液稀释细胞至 1.0×10^6 个细胞/ml。每升细胞培养物用 25 ml His-NAP1 病毒感染。

2. 在感染后 72 h，用临床离心机 4℃，2000 r/min 离心 5 min，收集细胞于 250 ml 锥形管中。冷 PBS 重悬每份细胞沉淀（原培养物体积的 1/10），重复离心。

 细胞沉淀可在液氮中冷冻后储藏在－80℃，在进一步处理前可保存数周。

3. 重悬细胞沉淀在裂解液 H 中（1/40 的原培养物体积）。用 Dounce 匀浆器“A”杵在 30 min 内打 40 次均质细胞。4℃，14 500*g*（SS-34 转头 11 000 r/min）离心 10 min，转移所有上清到一 50 ml 锥形管。

步骤 2：通过镍亲和层析纯化 NAP1

1. 按每 500 ml 原细胞培养体积，在裂解液 H 中平衡 1 ml Ni-NTA 琼脂糖树脂。加入细胞提取物并在颠倒旋转器中孵育 3～4 h。在临床离心机中 4℃，2000 r/min 离心 5 min 沉淀树脂。用 100 ml 裂解缓冲液洗涤树脂两次，然后用 100 ml 洗涤缓冲液 H 通过颠倒试管重悬树脂，在临床离心机中 2000 r/min 离心 3 min 沉淀树脂。

2. 洗脱蛋白质，通过轻轻涡旋将树脂重悬在 2 ml 洗脱缓冲液 H 中，在冰上孵育 5 min。在临床离心机中 2000 r/min 离心 3 min，然后将上清液转移到置于冰上的新管中。重复此洗脱循环三次，汇集洗脱液。

3. 在 12 000～15 000 MWCO 管中透析洗脱的 NAP1 两次，每次在 4 L 含 0.1 mol/L NaCl 的 HEGD 缓冲液中透析 2 h。

4. 在 4 L 含有 0.1 mol/L NaCl 的 NAP1 缓冲纯化中再透析 2 h。经透析的蛋白质可在液氮中冷冻，并在－80℃保存，以备进一步处理。

5. 在 15 ml 锥形管中 4℃，14 500*g*（SS34 转头 11 000 r/min）离心除去沉淀。用 SDS-PAGE 分析透析的 NAP1 和 BSA 标准物，估计蛋白质的量。

步骤 3：阴离子交换色谱纯化 NAP1

1. 使用 FPLC，根据制造商的指示用 15Q 树脂装填 HR-5 或 HR-b 柱。每 5 mg 在

步骤 2：5 中获得的 NAP1 使用 1 ml 包装树脂。在 10 倍柱体积的含 0.1 mol/L NaCl 的 NAP1 纯化缓冲液中平衡 Source 15Q 柱。

2. 加载 NAP1 到原 15Q 柱。用 10 倍柱体积的含 0.1 mol/L NaCl 的 NAP1 纯化缓冲液洗涤样品。用 20 倍柱体积的含 0.2～0.5 mol/L NaCl 的 NAP1 纯化缓冲液洗脱蛋白质。

NAP1 应洗脱在两个不同的峰中。早期（较低盐）峰抑制组装，而晚期（高盐）峰具有活性。早期峰洗脱一个 14 kDa 条带。收集 0.25～0.5 柱体积的馏分。

3. 每个馏分取 2 μl 进行 15% SDS-PAGE 凝胶电泳，鉴定纯的含 NAP1 的部分。合并峰馏分并透析两次，每次使用 2 L 含 0.1 mol/L NaCl 的 NAP1 纯化缓冲液透析 2 h。

4. 使用 8% 的 SDS-PAGE 凝胶同步电泳 NAP1 和 BSA 标准物，确定经透析的 NAP1 的浓度。将此材料 100～200 μl 分装在 1.5 ml 硅化试管并在液氮中冷冻，储存于 −80℃。

每升细胞培养物预期收获 1～3 mg 活性 NAP1。

疑难解答

重建后检测不到 DNA

可能原因：DNA、组蛋白或缓冲液中核酸酶污染。

解决方案：检测核酸酶，方法如下。将 DNA 与每个组分 37℃孵育 30 min，之后用琼脂糖凝胶电泳检查 DNA 的完整性核酸。透析缓冲液中 EDTA 的浓度也可提高到 1 mmol/L。

可能原因：重建过程中 DNA 聚集，导致沉淀。

解决方案：重新分析储存的 DNA 和组蛋白储存物的浓度，以确认组装的适当比例。因为自由的核苷酸可影响 DNA 的吸光度读数，琼脂糖凝胶电泳检测 DNA 模板分装样品，以确认吸光度读数。同样，用 SDS-PAGE 检查组蛋白储存物，并以牛血清白蛋白浓度为标准平行电泳。

可能原因：用与分析的 MNase 储存液浓度太高或终止缓冲液不正确。

解决方案：不用 MNase 消化，检查重建的完整性。对于阵列模板（约 2.5 kb），重建的阵列比游离 DNA 片段迁移稍快。

核小体发生重建但阵列亚饱和

可能原因：八聚体 DNA 或组蛋白浓度不准确。

解决方案：通过吸收法和琼脂糖凝胶或 SDS-PAGE 电泳重新检查浓度。

可能原因：八聚体储存液不能完全胜任组装。

解决方案：在某些情况下，重组八聚体储存液含有超过 10% 的无能组装的八聚体，这可能是由于不完全折叠所导致的。当使用含有各种氨基酸替换的八聚体时，这个比例可能还会增加。在这种情况下，可以简单地增加 r 值（如每 200 bp 重复使用 1.3～1.8 个八聚体），以实现更大的重建效率。

参考文献

Allfrey V. G., Faulkner R., and Mirsky A. E. 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci.* 51: 786-794.

An W. and Roeder R. G. 2004. Reconstitution and transcriptional analysis of chromatin in vitro. *Methods Enzymol.* 377: 460-474.

An W., Kim J., and Roeder R. G. 2004. Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. *Cell* 117: 735-748.

Angelov D., Verdel A., An W., Bondarenko V., Hans F., Doyen C. M., StuditskyV. M., Hamiche A., Roeder R. G., Bouvet P., and Dimitrov S. 2004. SWI/SNF remodeling and p300-dependent transcription of histone variant H2ABbd nucleosomal arrays. *EMBO J.* 23: 3815-3824.

Arents G., Burlingame R. W., Wang B. C., Love W. E., and Moudrianakis E. N. 1991. The nucleosomal core histone octamer at 3.1 Å resolution: A tripartite protein assembly and a left-handed superhelix. *Proc. Natl. Acad. Sci.* 88: 10148-10152.

Ausio J. 2000. Analytical ultracentrifugation and the characterization of chromatin structure. *Biophys. Chem.* 86: 141-153.

Balasubramanian R., Pray-Grant M. G., Selleck W., Grant P. A., and Tan S. 2002. Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. *J. Biol. Chem.* 277: 7989-7995.

Bao Y., Konesky K., Park Y. J., Simona R., Dyer P. N., Rangasamy D., Tremethick D. J., Laybourn P. J., and Luger K. 2004. Nucleosomes containing the histone variant H2A. Bbd organize only 118 base pairs of DNA. *EMBO J.* 23: 3314-3324.

Barton M. C., Madani N., and Emerson B. M. 1993. The erythroid protein cGATA-1 functions with a stage-specific factor to activate transcription of chromatin-assembled β-globin genes. *Genes Dev.* 7: 1796-1809.

Becker R. B., Tsukiyama T., and Wu C. 1994. Chromatin assembly extracts from Drosophila embryos. *Methods Cell Biol.* 44: 207-223.

Belmont A. S. and Bruce K. 1994. Visualization of G1 chromosomes: A folded, twisted, supercoiled chromonema model of interphase chromatid structure. *J. Cell Biol.* 127: 287-302.

Berndsen C. E., Albaugh B. N., Tan S., and Denu J. M. 2007. Catalytic mechanism of a MYST family histone acetyltransferase. *Biochemistry* 46: 623-629.

Blank T. A. and Becker R. B. 1995. Electrostatic mechanism of nucleosome spacing. *J. Mol. Biol.* 252: 305-313.

Boyer L. A., Logie C., Bonte E., Becker P. B., Wade P. A., Wolfe A. P., Wu C., Imbaizano A. N., and Peterson C. L. 2000. Functional delineation of three groups of the ATP-dependent family of chromatin remodeling enzymes. *J. Biol. Chem.* 275: 18864-18870.

Brownell J. E. and Allis C. D. 1995. An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in Tetrahymena macronuclei. *Proc. Natl. Acad. Sci.* 92: 6364-6368.

Brownell I. E., Zhou I., Ranalli T., Kobayashi R., Edmondson O. G., Roth S. Y., and Allis C. D. 1996. Tetrahymena histone acetyltransferase A: A homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84: 843-851.

Bruno M., Flaus A., Stockdale C., Rencurel C., Ferreira H., and Owen-Hughes T. 2003a. Histone H2A/h2B dimer exchange by ATP-dependent chromatin remodeling activities. *Mol. Cell* 12: 1599-1606.

Cairns B. R. 2005. Chromatin remodeling complexes: Strength in diversity, precision through specialization. *Curr. Opin. Genet. Dev.* 15: 185-190.

Cairns B. R. 2007. Chromatin remodeling: Insights and intrigue from single-molecule studies. *Nat. Struct. Mol. Biol.* 14: 989-996.

Cairns B. R., Kim Y. -J., Sayre M. H., Laurent B. C., and Kornber R. D. 1994. A multisubunit complex con-

taining the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 gene products isolated from yeast. *Proc. Natl. Acad. Sci.* 91: 1950-1954.

Carmen A. A., Rundlett S. E., and Grunstein M. 1996. HDA1 and HDA3 are components of a yeast histone deacetylase (HDA) complex. *J. Biol. Chem.* 271: 15837-15844.

Carruthers L. M. and Hansen J. C. 2000. The core histone N termini function independently of linker histones during chromatin condensation. *J. Biol. Chem.* 275: 37285-37290.

Carruthers L. M., Tse C., Walker K. P., 3rd, and Hansen I. C. 1999. Assembly of defined nucleosomal and chromatin arrays from pure components. *Methods Enzymol.* 304: 19-35.

Chadwick B. P. and Willard H. F. 2001a. A novel chomatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosoome. *J. Cell Blol.* 152: 375-384.

Chadwick B. P. and Willard H. F. 2001b. Histone H2A variants and the inactive X chromosome: Identification of a second macroH2A variant. *Hum. Mol. Genet.* 10: 1101-1113.

Chandy M., Gutierrez J. L., Prochasson P., and Workman J. L. 2006. SWI/SNF displaces SAGA-acetylated nucleosomes. Eukaryot. *Cell* 5: 1738-1747.

Chang B., Chen Y., Zhao Y., and Bruick R. K. 2007. JMJD6 is a histone arginine demethylase. *Science* 318: 444-447.

Chang C. H. and Lose D. S. 1997. The H3/H4 tetramer blocks transcript elongation by RNA polymerase II in vitro. *J. Biol. Chem.* 272: 23427-23434.

Cirillo L. A., Lin F. R., Cuesta I., Friedman I., Jarnik M., and Zaret K. S. 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* 9: 279-289.

Cote J., Quinn J., Workman J. L., and Peterson C. L. 1994. Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science* 265: 53-60.

Croston G. E., Lira L. M., and Kadonaga J. T. 1991. A general method for purification of H1 histones that are active for repression of basal RNA polymerase II transcription. *Protein Expr. Purif.* 2: 162-169.

Cuthbert G. L., Daujat S., Snowden A. W., Erdjument-Bromage H., Hagiwara T., Yamada M., Schneider R, Gregory P. D., Tempst P., Bannister A. J., and Kouzarides T. 2004. Histone deimination antagonizes arginine methylation. *Cell* 118: 545-553.

David R., Richter M. E, and Beck-Sickinger A. G. 2004. Expressed protein ligation. Method and applications. *Eur. J. Biochem.* 271: 663-677.

Dong F., Hansen J. C., and van Holde K. E. 1990. DNA and protein determinants of nucleosome positioning on sea urchin 5S rRNA gene sequences in vitro. *Proc. Natl. Acad. Sci.* 87: 5724-5728.

Dorigo B., Schaich T., Bystricky K., and Richmond T. J. 2003. Chromatin fiber folding: Requirement for the histone H4 N-terminal tail. *J. Mol. Biol.* 327: 85-96.

Durr H., Flaus A., Owen-Hughes T., and Hopfner K. P. 2006. Snf2 family ATPases and DExx box helicases: Differences and unifying concepts from high-resolution crystal structures. *Nucleic Acids Res.* 34: 4160-4167.

Eickbush T. H. and Moudrianakis E. N. 1978. The histone core complex: An octamer assembled by two Sets of protein-protein interactionS. *Biochemistry* 17: 4955-4964.

Esteve P. O., Chin H. G., Smallwood A., Feehery G. R., Gangisetty O., Karpf A. R., Carey M. F., and Pradhan S. 2006. Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes Dev.* 20: 3089-3103.

Fan Y., Nikitina T., Morin-Kensicki EM., Zhao J., Magnuson T. R., Woodcock C. L., and Skoultchi A. I. 2003. H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo. *Mol. Cell Biol.* 23: 4559-4572.

Fan Y., Nikitina T., Zhao J., Fleury T. J., Bhattacharyya R., Bouhassira E. E., Stein A., Woodcock C. L., and Skoultchi A. I. 2005. Histone H1 depletion in mammals alters global chromatin structure but causes specific

changes in gene regulation. *Cell* 123: 1199-1212.

Fang J., Feng Q., Ketel C. S., Wang H., Cao R., Ala L., ErdjumentBromage H., Tempst P., Simon J. A., and Zhang Y. 2002. Purification and functional characterization of SET8, a nucleosomal histone H4-lysine 20-specific methyltransferase. *Curr. Biol.* 12: 1086-1099.

Ferreira H., Flaus A., and Owen-Hughes T. 2007. Histone modifications influence the action of Snf2 family remodelling enzymes by different mechanisms. *J. Mol. Biol.* 374: 563-579.

Flaus A., Rencurel C., Ferreira H., Wiechens N., and Owen-Hughes T. 2004. Sin mutations alter inherent nucleosome mobility. *EMBO J.* 23: 343-353.

Fletcher T. M. and Hansen J. C. 1996. The nucleosomal array: Structure/function relationships. *Crit. Rev. Eukaryot. Gene Expr.* 6: 149-188.

Fry C. J., Shogren-Knaak M. A., and Peterson C. L. 2004. Histone H3 amino-terminal tail phosphorylation and acetylation: Synergistic or independent transcriptional regulatory marks? *Cold Spring Harbor Symp. Quant. Biol.* 69: 219-226.

Fyodorov D. V. and Kadonaga J. T. 2003. Chromatin assembly in vitro with purified recombinant ACF and NAP-1. *Methods Enzymol.* 371: 499-515.

Garcia B. A., Hake S. B., Diaz R. L., Kauer M., Morris S. A., Recht J., Shabanowitz J., Mishra N., Strahl B. D., Allis C. D., and Hunt D. F. 2007. Organismal differences in post-translational modifications in histones H3 and H4. *J. Biol. Chem.* 282: 7641-7655.

Garcia-Ramirez M., Leuba S., and Ausio J. 1990. One-step fractionation method for isolation H1 histones from chromatin under nondenaturing conditions. *Protein Expr. Purif.* 1: 40-44.

Georgakopoulos T. and Thireos G. 1992. Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. *EMBO J.* 11: 4145-4152.

Glikin G. C., Ruberti I., and Worcel A. 1984. Chromatin assembly in Xenopus oocytes: In vitro studies. *Cell* 37: 33-41.

Gloss L. M. and Placek B. J. 2002. The effect of salts on the stability of the H2A/H2B histone dimer. *Biochemistry* 41: 14951-14599.

Godde J. S. and Wolffe A. P. 1995. Disruption of reconstituted nucleosomes. The effect of particle concentration, $MgCl_2$ and KCI concentration, the histone tails, and temperature. *J. Biol. Chem.* 270: 27399-27402.

Guermah M., Palhan V. B., Tackett A. I., Chait B. T., and Roeder R. G. 2006. Synergistic functions of SII and p300 in productive activator-dependent transcription of chromatin templates. *Cell* 125: 275-286.

Gutierrez J. L., Chandy M., Carrozza M. J., and Workman J. L. 2007. Activation domains drive nucleosome eviction by SWI/SNF. *EMBO J.* 26: 730-740.

Hamiche A., Sandaltzopoulos R., Gdula D. A., and Wu C. 1999. ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. *Cell* 97: 833-842.

Hansen J. C. and Lohr D. 1993. Assembly and structural properties of subsaturated chromatin arrays. *J. Biol. Chem.* 268: 5840-5848.

Hansen J. C. and Wolfe A. P. 1994. A role for histones H2A/H2B in chromatin folding and transcriptional repression. *Proc. Natl. Acad. Sci.* 91: 2339-2343.

Hansen J. C., Lebowitz J., and Demeler B. 1994. Analytical ultracentrifugation of complex macromolecular systems. *Biochemistry* 33: 13155-13163.

Hansen J. C., van Holde K. E., and Lohr D. 1991. The mechanism of nucleosome assembly onto oligomers of the sea urchin 5 S DNA positioning sequence. *J. Biol. Chem.* 266: 4276-4282.

Hansen J. C., Ausio J., Stanik V. H., and van Holde K. E. 1989. Homogeneous reconstituted oligonucleosomes, evidence for salt-dependent folding in the absence of histone Hl. *Biochemistry* 28: 9129-9136.

Hansen J. C., Kreider J. I., Demeler B., and Fletcher T. M. 1997. Analytical ultracentrifugation and agarose gel

electrophoresis as tools for studying chromatin folding in solution. *Methods* 12: 62-72.

Hassan A. H., Neely K. E., and Workman J. L. 2001. Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. *Cell* 104: 817-827.

Hayes J. J., Tullius T. D., and Wolffe A. P. 1990. The structure of DNA in a nucleosome. *Proc. Natl. Acad. Sci.* 87: 7405-7409.

He S., Bauman D., Davis J. S., Loyola A., Nishioka K., Gronlund J. L., Reinberg D., Meng F., Kelleher N., and McCafferty D. G. 2003. Facile synthesis of site-specifically acetylated and methylated his-tone proteins: Reagents for evaluation of the histone code hypothesis. *Proc. Natl. Acad. Sci.* 100: 12033-12038.

Horn P. J., Crowley K. A., Carruthers L. M., Hansen I. C., and Peterson C. L. 2002. The SIN domain of the histone octamer is essential for intramolecular folding of nucleosomal arrays. *Nat. Struct. Biol.* 9: 167-171.

Huyen Y., Zgheib O., Ditullio R. A., Jr., Gorgoulis V. G., Zacharatos P., Petty T. J., Sheston E. A., Mellert H. S., Stavridi E. S., and Halazonetis T. D. 2004. Methylated lysine 79 of histone H3 targets 53BP ito DNA double-strand breaks. *Nature* 432: 406-411.

Ikeda K., Steger D. J., Eberharter A., and Workman J. L. 1999. Activation domain-specific and general transcription stimulation by native histone acetyltransferase complexes. *Mol. Cell Biol.* 19: 855-863.

Imai S., Armstrong C. M., Kaeberlein M., and Guarente L. 2000. Transcriptional silencing and longevity protein Sir2 is an NADdependent histone deacetylase. *Nature* 403: 795-800.

Imbalzano, A. N., Schnitzler, G. R., and Kingston R. E. 1996. Nucleosome disruption by human SWI/SNF is maintained in the absence of continued ATP hydrolysis. *J. Biol. Chem.* 271: 20726-20733.

Imbalzano A. N., Kwon H., Green M. R., and Kingston R. E. 1994. Facilitated binding of TATA-binding protein to nucleosomal DNA [see comments]. *Nature* 370: 481-485.

Inoue A. and Fujimoto D. 1969. Enzymatic deacetylation of histone. *Biochem. Biophys. Res. Commun.* 36: 146-150.

Ishimi Y. and Kikuchi A. 1991. Identification and molecular cloning of yeast homolog of nucieosome assembly protein I which facilitates nucleosome assembly in vitro. *J. Biol. Chem.* 266: 7025-7029.

Ito T., Bulger M., Kobayashi R., and Kadonaga J. T. 1996. Drosophila NAP-1 is a core histone chaperone that functions in ATP-facilitated assembly of regularly spaced nucleosomal arrays. *Mol. Cell Biol.* 16: 3112-3124.

Ito T., Levenstein M. E., Fyodorov D. V., Kutach A. K., Kobayashi R., and Kadonaga J. T. 1999. ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. *Genes Dev.* 13: 1529-1539.

Jair K. W., Bachman K. E., Suzuki H., Ting A. H., Rhee I., Yen R. W., Baylin S. B., and Schuebel K. E. 2006. De novo CpG island methylation in human cancer cells. *Cancer Res.* 66: 682-692.

Jiang W., Nordeen S. K., and Kadonaga J. T. 2000. Transcriptional analysis of chromatin assembled with purified ACF and dNAP1 reveals that acetyl-CoA is required for preinitiation complex assembly. *J. Biol. Chem.* 275: 39819-39822.

Juan L. J., Utley R. T., Adams C. C., Vettese-Dadey M., and Workman J. L. 1994. Differential repression of transcription factor binding by histone Hl is regulated by the core histone amino termini. *EMBO J.* 13: 6031-6040.

Kamakaka R. T. and Kadonaga J. T. 1994. The soluble nuclear fraction, a highly efficient transcription extract from Drosophila embryos. *Methods Cell Biol.* 44: 225-235.

Kan P. Y. and Hayes J. J. 2007. Detection of interactions between nudeosome arrays mediated by specific core histone tail domains. *Methods* 41: 278-285.

Kassabov S. R., Zhang B., Persinger J., and Bartholomew B. 2003. SWI/SNF unwraps, slides, and rewraps the nucleosome. *Mol. Cell* 11: 391-403.

Kilgore J. A., Hoose S. A., Gustafson T. L., Porter W., and Kladde M. P. 2007. Single-molecule and popula-

tion probing of chromatin structure using DNA methyltransferases. *Methods* 41: 320-332.

Klose R. J. and Bird A. P. 2006. Genomic DNA methylation: The mark and its mediators. *Trends Biochem. Sci.* 31: 89-97.

Klose R. J. and Zhang Y. 2007. Regulation of histone methylation by demethylimination and demethylation. *Nat. Rev. Mol. Cell Biol.* 8: 307-318.

Kornberg R. D. 1974. Chromatin structure: A repeating unit of his-tones and DNA. *Science* 184: 868-871.

Kornberg R. D. and Thomas J. O. 1974. Chromatin structure; oligomers of the histones. *Science* 184: 865-868.

Kouzarides T. 2007. Chromatin modifications and their function. *Cell* 128: 693-705.

Kruger W., Peterson C. L., Sil A., Coburn C., Arents G., Moudrianakis E. N., and Herskowitz I. 1995. Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription. *Genes Dev.* 9: 2770-2779.

Kuzmichev A., Nishioka K., Erdjument-Bromage H., Tempat P., and Reinberg D. 2002. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* 16: 2893-2905.

Kwon H., Imbaizano A. N., Khavari P. A., Kingston R. E., and Green M. R. 1994. Nucleosome disruption and enhancement of activator binding by a human SWI/SNF complex [see comments]. *Nature* 370: 477-481.

Langer M. R., Tanner K. G., and Denu J. M. 2001. Mutational analysis of conserved residues in the gcn5 family of histone acetyltransferases. *J. Biol. Chem.* 276: 31321-31331.

Laskey R. A., Mills A. D., and Morris N. R. 1977. Assembly of SV4O chromatin in a cell-free system from Xenopus eggs. *Cell* 10: 237-243.

Lee K. M., Chafin D. R., and Hayes J. J. 1999. Targeted cross-linking and DNA cleavage within model chromatin complexes. *Methods Enzymol.* 304: 231-251.

Logie C. and Peterson C. L. 1997. Catalytic activity of the yeast SWI/SNF complex on reconstituted nucleosome arrays. *EMBO J.* 16: 6772-6782.

Logie C. and Peterson C. L. 1999. Purification and biochemical properties of yeast SWI/SNF complex. *Methods Enzymol.* 304: 726-741.

Lorch Y., Maier-Davis B., and Kornberg R. D. 2006. Chromatin remodeling by nucleosome disassembly in vitro. *Proc. Natl. Acad. Sci.* 103: 3090-3093.

Lorch Y., Zhang M., and Kornberg R. D. 1999. Histone octamer transfer by a chromatin-remodeling complex. *Cell* 96: 389-392.

Lowary P. T. and Widom J. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nudeosome positioning. *J. Mol. Biol.* 276: 19-42.

Luger K., Rechsteiner T. J., and Richmond T. J. 1999. Expression and purification of recombinant histones and nucleosome reconstitution. *Methods Mol. Biol.* 119: 1-16.

Luger K., Mader A. W., Richmond R. K., Sargent D. F., and Richmond T. J. 1997. Crystal structure of the nucleosome core particle at 2. 8 A resolution. *Nature* 389: 251-260.

Luk E., Vu N. D., Patteson K., Mizuguchi G., Wu W. H., Ranjan A., Backus J., Sen S., Lewis M., Bai Y., and Wu C. 2007. Chz1, a nuclear chaperone for histone H2AZ. *Mol. Cell* 25: 357-368.

Malik S., Wallberg A. E., Kang Y. K., and Roeder R. G. 2002. TRAP/SMCC/mediator-dependent transcriptional activation from DNA and chromatin templates by orphan nuclear receptor hepatocyte nuclear factor 4. Mol. *Cell Biol.* 22: 5626-5637.

Marcus G. A., Silverman N., Berger S. L., Horiuchi J., and Guarente L. 1994. Functional similarity and physical association between GCN5 and ADA2: Putative transcriptional adaptors. *EMBO J.* 13: 4807-4815.

Marino-Ramirez L., Hsu B., Baxevanis AD., and Landsman D. 2006. The Histone Database: A comprehensive resource for histones and histone fold-containing proteins. *Proteins* 62: 838-842.

Mavrich T. N., Jiang C., Ioshikhes I. P., Li X. Y., Venters B. J., Zanton S. J., Tomsho L. P., Qi J., Glaser R. L., Schuster S. C., Gilmour D. S., Albert I., and Pugh B. F. 2008. Nucleosome organization in the Drosophila genome. *Nature* 453: 358-362.

Mizuguchi G., Shen X., Landry J., Wu W. H., SenS., and Wu C. 2003. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* 303: 343-348.

Mizzen C. A., Brownell I. E., Cook R. G., and Allis C. D. 1999. Histone acetyltransferases: preparation of substrates and assay procedures. *Methods Enzymol*. 304: 675-696.

Moehs C. R, Baxevanis A. D., Moudrianakis E. N., and Spiker S. 1992. Enhanced stability of histone octamers from plant nucleosomes: Role of H2A and H2B histones. *Biochemistry* 31: 10844-10851.

Neely K. E., Hassan A. H., Wallberg A. E., Steger D. J., Cairns B. R., Wright A. P., and Workman J. L. 1999. Activation domain-mediated targeting of the SWI/SNF complex to promoters stimulates transcription from nucleosome arrays. *Mol. Cell* 4: 649-655.

Oudet P., Gross-Bellard M., and Chambon P. 1975. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* 4: 281-300.

Owen-Hughes T. and Workman J. L. 1996. Remodeling the chromatin structure of a nucleosome array by transcription factor-targeted trans-displacement of histones. *EMBO J*. 15: 4702-4712.

Owen-Hughes T., Utley R. T., Cote J., Peterson C. L., and Workman J. L. 1996. Persistent site-specific remodeling of a nucleosome array by transient action of the SWI/SNF complex. *Science* 273: 513-516.

Owen-Hughes T., Utley R. T., Steger D. J., West J. M., John S., Cote J., Havas K. M., and Workman J. L. 1999. Analysis of nucleosome disruption by ATP-driven chromatin remodeling complexes. *Methods Mol. Biol.* 119: 319-331.

Park Y. J., Chodaparambil J. V., Bao V., McBryant S. J., and Luger K. 2005. Nucleosome assembly protein 1 exchanges histone H2A/H2B dimers and assists nucleosome sliding. *J. Biol. Chem.* 280: 1817-1825.

Pazin M. I., Hermann J. W., and Kadonaga J. T. 1998. Promoter structure and transcriptional activation with chromatin templates assembled in vitro. A single Gal4-VP16 dimer binds to chromatin or to DNA with comparable affinity. *J. Biol. Chem.* 273: 34653-34660.

Pazin M. J., Bhargava P., Geiduschek E. P., and Kadonaga J. T. 1997. Nucleosome mobility and the maintenance of nucleosome positioning. *Science* 276: 809-812.

Pederson D. S. and Fidrych T. 1994. Heat shock factor can activate transcription while bound to nucleosomal DNA in Saccharomyces cerevisiae. *Mol. Cell Biol*. 14: 189-199.

Phillips D. M. and Johns E. W. 1965. A fractionation of the histones of group F2a from calf thymus. *Biochem. J.* 94: 127-130.

Polach K. J. and Widom J. 1995. Mechanism of protein access to specific DNA sequences in chromatin: A dynamic equilibrium model for gene regulation. *J. Mol. Biol.* 254: 130-149.

Pollard K. J., Samuels M. L., Crowley K. A., Hansen J. C., and Peterson C. L. 1999. Functional interaction between GCN5 and polyamines: A new role for core histone acetylation. *EMBO J*. 18: 5622-5633.

Pradhan S., Bacolla A., Wells R. D., and Roberts R. J. 1999. Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. *J. Biol. Chem.* 274: 33002-33010.

Raisner R. M., Hartley P. D., Meneghini M. D., Baa M. Z., Liu C. L., Schreiber S. L., Rando O. J., and Madhani H. D. 2005. Histone variant H2A. Z marks the 5′ ends of both active and inactive genes in euchromatin. *Cell* 123: 233-248.

Rippe K., Schrader A., Riede R, Strohner R., Lehmann E., and Langst G. 2007. DNA sequence-and conformation-directed positioning of nucleosomes by chromatin-remodeling complexes. *Proc. Natl. Acad. Sci.* 104: 15635-15640.

Roberts M. S., Fragoso G., and Hager G. L. 1995. Nucleosomes reconstituted in vitro on mouse mammary tumor virus B region DNA occupy multiple translational and rotational frames. *Biochemistry* 34: 12470-12480.

Rountree M. R., Bachman K. E., and Baylin S. B. 2000. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat. Genet.* 25: 269-277.

Schones D. E., Cui K., Cuddapah S., Rob T. Y., Barski A., Wang Z., Wei G., and Zhao K. 2008. Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132: 887-898.

Shen X., Ranallo R., Choi E., and Wu C. 2003. Involvement of actinrelated proteins in ATP-dependent chromatin remodeling. *Mol. Cell* 12: 147-155.

Shi Y., Lan F., Matson C., Mulligan P., Whetstine J. R., Cole R. A., Casero R. A., and Shi Y. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119: 941-953.

Shogren-Knaak M. A. and Peterson C. L. 2004. Creating designer his-tones by native chemical ligation. *Methods Enzymol.* 375: 62-76.

Shogren-Knaak M. A., Fry C. J., and Peterson C. L. 2003. A native peptide ligation strategy for deciphering nucleosomal histone modifications. *J. Biol. Chem.* 278: 15744-15748.

Shogren-Knaak M., Ishii H., Sun J. M., Pazin M. J., Davie J. R., and Peterson C. L. 2006. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311: 844-847.

Shrader T. E. and Crothers D. M. 1989. Artificial nucleosome positioning sequences. *Proc. Natl. Acad. Sci.* 86: 7418-7422.

Shundrovsky A., Smith C. L., Lis J. T., Peterson C. L., and Wang M. D. 2006. Probing SWI/SNF remodeling of the nucleosome by unzipping single DNA molecules. *Nat. Struct. Mol. Biol.* 13: 549-554.

Simon M. D., Chu F., Racki L. R., de Ia Cruz C. C., Burlingame A. L., Panning B., Narlikar G. J., and Shokat K. M. 2007. The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* 128: 1003-1012.

Simon R. H. and Felsenfeld G. 1979. A new procedure for purifying histone pairs H2A + H2B and H3 + H4 from chromatin using hydroxylapatite. *Nucleic Acids Res.* 6: 689-696.

Simpson R. T., Thoma F., and Brubaker J. M. 1985. Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: A model system for study of higher order structure. *Cell* 42: 799-808.

Sinha M. and Peterson C. L. 2008. A Rad51 presynaptic filament is sufficient to capture nucleosomal homology during recombinational repair of a DNA double-strand break. *Mol. Cell* 30: 803-810.

Smith C. L. and Peterson C. L. 2005. ATP-dependent chromatin remodeling. *Curr. Top Dev. Biol.* 65: 115-148.

Steger D. J. and Workman J. L. 1997. Stable co-occupancy of transcription factors and histones at the HIV-1 enhancer. *EMBO J.* 16: 2463-2472.

Steger D. J., Eberharter A., John S., Grant P. A., and Workman J. L. 1998. Purified histone acetyltransferase complexes stimulate HIV-1 transcription from preassembled nucleosomal arrays. *Proc. Natl. Acad. Sci.* 95: 12924-12929.

Stein A. and Mitchell M. 1988. Generation of different nucleosome spacing periodicities in vitro. Possible origin of cell type specificity. *J. Mol. Biol.* 203: 1029-1043.

Taunton J., Hassig C. A., and Schreiber S. L. 1996. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p [see comments]. *Science* 272: 408-411.

Terrell A. R., Wongwisansri S., Pilon J. L., and Laybourn P. J. 2002. Reconstitution of nucleosome positioning, remodeling, histone acetylation, and transcriptional activation on the PHO5 promoter. *J. Biol. Chem.* 277: 31038-31047.

Trojer P., Li, G., Sims III R. J., Vaquero A., Kalakonda N., Boccuni P., Lee, D., Erdjument-Bromage H., Tempst P., Nimer S. D., Wang Y. H., and Reinberg D. 2007. L3MBTL1, a histone-methylation dependent chromatin lock. *Cell* 129: 915-928.

Tse C.，Sera T.，Wolffe A. P.，and Hansen J. C. 1998. Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. *Mol. Cell. Biol.* 18：4629-4638.

Tsukada Y.，Fang J.，Erdjument-Bromage H.，Warren M. E.，Borchers C. H.，Tempst P.，and Zhang Y. 2006. Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439：811-816.

Usachenko S. I.，Bavykin S. G.，Gavin I. M.，and Bradbury E. M. 1994. Rearrangement of the histone H2A Cterminal domain in the nucleosome. *Proc. Natl. Acad. Sci.* 91：6845-6849.

Utley R. T.，Ikeda K.，Grant P. A.，Cote J.，Steger D. J.，Eberharter A.，John S.，and Workman J. L. 1998. Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. *Nature* 394：498-502.

Vary J. C.，Fazzio T. G.，and Tsukiyama T. 2004. Assembly of yeast chromatin using ISWI complexes. *Methods Enzymol.* 375：88-102.

Vaziri H.，Dessain S. K.，Ng Eaton E.，Imai S. I.，Frye R. A.，Pandita T. K.，Guarente L.，and Weinberg R. A. 2001. hSIR2（SIRT1）functions as an NAD-dependent p53 deacetylase. *Cell* 107：149-159.

Vettese-Dadey M.，Walter P.，Chen H.，Juan L. J.，and Workman J. L. 1994. Role of the histone amino termini in facilitated binding of a transcription factor，GAL4-AH，to nucleosome cores. *Mol. Cell. Biol.* 14：970-981.

Vettese-Dadey M.，Grant P. A.，Hebbes T. R.，Crane-Robinson C.，Allis C. D.，and Workman J. L 1996. Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nudeosomal DNA in vitro. *EMBO J.* 15：2508-2518.

Vignali M.，Steger D. J.，Neely K. E.，and Workman J. L. 2000. Distribution of acetylated hiatones resulting from Ga14-VP16 recruitment of SAGA and NuA4 complexes. *EMBO J.* 19（11）：2629-2640.

Wade P. A.，Jones P. L.，Vermaak D.，and Wolffe A. P. 1999. Purification of a histone deacetylase complex from Xenopus laevis：Preparation of substrates and assay procedures. *Methods Enzymol.* 304：715-725.

Wailberg A. E.，Pedersen K.，Lendahl U.，and Roeder R. G. 2002. p300 and PCAF act cooperatively to mediate transcriptional activation from chromatin templates by notch intracellular domains in vitro. *Mol. Cell Biol.* 22：7812-7819.

Weintraub H. and Groudine M. 1976. Chromosomal subunits in active genes have an altered conformation. *Science* 193：848-856.

Whetstine J. R.，Notthe A.，Lan F.，Huarte M.，Smolikov S.，Chen Z.，Spooner E.，Li E.，Zhang G.，Colaiacovo M.，and Shi Y. 2006. Reversal of histone lysine trimethylat ion by the JMJD2 family of histone demethylases. *Cell* 125：467-481.

Wisniewski J. R.，Zougman A.，Kruger S.，and Mann M. 2007. Mass spectrometric mapping of linker histone Hi variants reveals multiple acetylations，methylations，and phosphorylation as well as differences between cell culture and tissue. *Mol. Cell. Proteomics* 6：72-87.

Wongwisansri S. and Laybourn P. I. 2004. Reconstitution of yeast chromatin using yNap1p. *Methods Enzymol.* 375：103-117.

Woodcock C. L.，Skoultchi A. I.，and Fan Y. 2006. Role of linker histone in chromatin structure and function：H1 stoichiometry and nucleosome repeat length. *Chromosome Res.* 14：17-25.

Workman J. L.，Taylor I. C.，Kingston R. E.，and Roeder R. G. 1991. Control of class II gene transcription during in vitro nucleosome assembly. *Methods Cell Biol.* 35：419-447.

Yang X.，Zaurin R.，Beato M.，and Peterson C. L. 2007. Swi3p controls SWI/SNF assembly and ATP-dependent H2A/H2B displacement. *Nat. Struct. Mol. Biol.* 14：540-547.

Yudkovsky N.，Logie C.，Hahn S.，and Peterson C. L. 1999. Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators. *Genes Dev.* 13：2369-2374.

Zhang H.，Roberts D. N.，and Cairns B. R. 2005. Genome-wide dynamics of Htzi，a histone H2A variant that poi-

ses repressed/basal promoters for activation through histone loss. *Cell* 123：219-231.

Zhang L.，Eugeni E. E.，Parthun M. R.，and Freitas M. A. 2003. Identification of novel histone post-translational modifications by peptide mass fingerprinting. *Chromosoma* 112：77-86.

Zhou J.，Fan J. Y.，Rangasamy D.，and Tremethick D. J. 2007. The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. *Nat. Struct. Mol. Biol.* 14：1070-1076.

附录：注 意 事 项

一般注意事项

请注意，本手册附录中的注意事项并不全面。读者应该始终向个别厂商和其他资源咨询当前的和特定的产品信息。在正文部分中讨论的化学试剂和其他材料不能根据符号〈!〉辨认，该符号仅在方案中表明危险材料。然而，如无特殊处理它们对使用者可能是危险的。更多信息请咨询当地的安全办公室或制造商的安全指导方针。

应时刻留心下列一般注意事项。

• 开始程序之前完全熟悉所使用物质的属性。

• 无警告并不一定意味着某物品是安全的，因为信息可能并不总是完整的和能得到的。

• 如果暴露于有毒物质，请立即与你当地的安全办公室联系，以获得指导。

• 对于所有化学的、生物的和放射性的废物都应使用适当的处理方法。

• 关于戴手套的特定使用说明，请咨询当地的安全办公室。

• 操作浓缩的酸和碱时要非常小心，要戴护目镜和合适的手套。当对大量酸碱进行操作时应戴面具。

不要将强酸与有机溶剂相混合，因为它们会发生反应。根据经验，特别是硫磺酸和硝酸会反应，并释放大量热量，导致失火和爆炸。

不要将强碱与卤化溶剂相混合，因为它们会形成导致爆炸的卡宾。

• 谨慎处理及储存压缩气体容器，因为它们可能含有易燃、有毒、腐蚀性气体，以及引起窒息的物质或氧化剂。关于恰当的程序，咨询必须由供应商提供的材料安全数据表。

• 绝不要用嘴吸移溶液，这种方法不卫生，且危险，应始终使用移液器或吸耳球。

• 将卤化溶剂和非卤化溶剂始终分开保存（如氯仿和丙酮混合，在碱存在的条件下会引起意外的反应）。卤化溶剂是有机溶剂，如氯仿、二氯甲、三氯三氟代乙烷及二氯乙烷。有些非卤化溶剂是戊烷、庚烷、乙醇、甲醇、苯、甲苯、二甲基甲酰胺（DMF）、二甲基亚砜（DMSO）及乙腈。

• 激光辐射，不管是可见的还是不可见的，都能造成对眼睛和皮肤的严重损害。采取适当的预防措施，以防止暴露于直接的和反射的光束。始终遵循制造商的安全指引，并咨询当地的安全办公室。更详细的信息请看下面的警告。

• 闪光灯，因为它们的光强度，可对眼睛有害。它们有时也可能会爆炸。穿戴合适的眼睛保护，并遵从制造商的指导方针。

• 摄影固定剂、显影剂和光阻材料也含有害的化学药品。应按照生产厂家的说明对它们进行小心操作。

• 电源和电泳设备如使用不当能引起严重的火灾和触电危险。

• 实验室中微波炉和高压灭菌锅的使用需要一定的预防措施。曾发生过与其使用相关的事故（如熔化保存在瓶中的琼脂或细菌培养用琼脂或灭菌）。如果螺旋盖没有完全取下，并且没有足够的空间使蒸汽排出，当把容器从微波炉或高压灭菌锅取出时，瓶子会发生爆炸，造成严重的损伤。在微波和高压灭菌之前，始终要完全移走瓶盖。对于常规的不需要灭菌的琼脂糖凝胶来说，一个替代方法是称取琼脂，并将溶液置于烧瓶中。

• 超声发生器用高频声波（16～100 kHz）破碎细胞和用于其他用途。这种“超声波”，通过空气传导，不会对人构成直接危害，但相关的高容量可听声会引起多种影响，包括头痛、恶心和耳鸣。应避免身体与高强度超声（非医学成像设备）的直接接触。使用恰当的耳朵保护，并且在使用该类设备的实验室门口设置标示。

• 操作诸如显微镜切片机刀刃、外科手术刀、剃刀、折页顶针等切裁装置时，要特别小心。显微镜切片机刀刃极其锋利，切片时要小心。如果不熟悉它们的使用，应请人演示正确程序。为了正确处置，在你的实验室使用“利器”处理容器。仍然附带注射筒，无防护丢弃用过的针头。这样做防止操作用过的针头时造成伤害（并且可能感染，参见生物安全），因为在设法更换针头防护套时，会发生很多事故。破碎的巴斯德吸液管、盖玻片或幻灯片也可能造成伤害。

• 动物待遇：任何时候都必须遵守对动物的人道待遇程序。咨询当地的动物机构，获得指导。动物（如老鼠）若已诱发过敏，则反复接触会增加这些过敏的强度。在操作动物时，始终穿戴实验服和手套。如果已知对头皮屑或唾液过敏，应戴上面罩。

常见化学药品的一般特性

危险材料名单可归纳如下类别。

• 无机酸，如盐酸、硫酸、硝酸或磷酸，是有刺痛蒸汽的无色液体。避免泼溅到皮肤或衣服上。若发生泼溅，应该用大量的水稀释。这些酸的浓缩形式可以摧毁纸、纺织品和皮肤，并可对眼睛造成严重伤害。

• 无机碱（如氢氧化钠）为白色固体，溶于水并产热。浓溶液会缓慢溶解皮肤，甚至指甲。

• 重金属盐通常是有色粉状固体，溶于水，其中许多是强有力的酶抑制剂，因此对人体和对环境（如鱼类和藻类）有毒性。

• 多数有机溶剂是易燃挥发液体。吸入蒸汽，会引起恶心或头晕，应予以避免，还要避免皮肤接触。

• 其他有机化合物，包括有机硫化合物（如巯基乙醇或有机胺），会有非常不愉快的气味。其他的则具有高度活性，应适当谨慎处理。

• 如果处理不当，染料及其溶液不仅可以染色样品，而且可以染色皮肤和衣物，它们中有些还是突变诱导物（如溴化乙锭）、致癌物和毒物。

• 几乎所有以“酶”结尾的名称［如过氧化氢酶、β-葡萄糖醛酸酶或裂解酶（zymolyase)］都是指酶。也有一些非系统命名的酶，如胃蛋白酶。其中许多酶由生产商以含有缓冲物质等的制剂提供，要知道包含在这些物质中的单个材料的性能。

• 有毒化合物通常用于操纵细胞，这些化合物是有危险的，应妥善处理。

• 请注意，所列出的几个化合物尚未彻底研究清楚它们的毒物学特性。处理每一种化学药品时要适当注意。虽然化合物的毒性作用是可以量化的（如 LD_{50}值），但致癌物和突变剂不可能接触一次就能产生效果。还要认识到与特定化合物相关的危险也可能取决于它的物理状态（细粉末与大晶体、乙醚与甘油、干冰与压力下毒气弹中的二氧化碳）。预计实验接触过程中哪种情况最有可能发生，并且怎样最好地保护自己和环境。

危险材料

注：一般情况下，专卖材料这里没有列出。试剂盒和其他商业名目，以及大多数麻醉剂、染料、固定剂和染色剂也没有包括在内。麻醉剂和抗生素也需要特别注意。遵照制造商的伴随这些产品所提供的安全指引。

乙酸（浓缩的）操作应特别小心。通过吸入、摄入或皮肤吸收，可能对人体有害。戴适当的手套和护目镜，在通风橱中使用。

乙酸（冰）具有很强的腐蚀性，必须格外小心地操作，它可能是一种致癌物。液体和雾气造成对所有机体组织严重的烧伤。通过吸入、摄入或皮肤吸收，可能对人体有害。戴适当的手套和护目镜，在通风橱中使用，远离热源、火花和明火。

丙烯酰胺（未聚合的）是一种很强的神经毒素，并可通过皮肤吸收（其作用具有累积性）。避免吸入其粉尘。当称量丙烯酰胺和亚甲基双丙烯酰胺时，戴合适的手套和面具，在化学通风橱中使用。一般认为聚丙烯酰胺无毒，但也应谨慎操作，因为它还可能会含有少量未聚合的丙烯酰胺。

放射菌素 D 是致畸剂和致癌剂。如果吸入、摄入或经皮肤吸收后具有强毒性，可能有致命危险，它还会引起刺激。避免吸入粉尘，戴合适的手套和安全眼镜，并始终在化学通风橱中使用。放线菌素 D 溶液对光敏感。

α-鹅膏蕈碱通过吸入、摄入或经皮肤吸收，具有强毒性，可能有致命危险。症状可能延迟长达 6～24 h，戴合适的手套和安全眼镜，并始终在化学通风橱中操作。

硫酸铁铵，$FeNH_4(SO_4)_2 \cdot 12H_2O$，通过吸入、摄入或经皮肤吸收，对人体有害。戴合适的手套和安全眼镜。

过硫酸铵，$(NH4)_2S_2O_8$，对黏膜和上呼吸道组织、眼睛和皮肤有极强的破坏性。戴合适的手套和安全眼镜，穿防毒服装。仅在化学通风橱中使用。操作后彻底洗手。

抑酶肽，通过吸入、摄入或经皮肤吸收，可能有害，也可能会引起过敏反应。接触后可能引起胃肠效应、肌肉疼痛、血压改变，或支气管痉挛。戴合适的手套和安全眼镜，不要吸入粉尘。

硼酸，H_3BO_3，通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜。

Bradford 染料，含有磷酸和甲醇，具有腐蚀性和毒性。戴合适的手套和安全眼镜。

溴酚蓝，通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜，并在化学通风橱中使用。

^{14}C，见放射性物质

$CaCl_2$，见氯化钙

氯化钙，$CaCl_2$，具有吸湿性并可能导致心脏紊乱。通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜。

二氧化碳，CO_2，不管是哪种形式，如通过吸入、摄入或经皮肤吸收，都可能是致命的。在高浓度时，可麻痹呼吸中枢，引起窒息。仅在通风良好的地方使用。当以干冰形式存在时，接触二氧化碳也能引起冻伤。不要将大量的干冰放在密闭空间，如冷藏室。戴合适的手套和护目镜。

CBR，见考马斯亮蓝

CH_3CH_2OH，见乙醇

$CHCl_3$，见氯仿

$C_7H_7FO_2S$，见苯甲基磺酰氟

氯霉素，是一种可能的致癌物，并且通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜，并在化学通风橱中使用。

氯仿，$CHCl_3$，对皮肤、眼睛、黏膜和呼吸道有刺激，是一种致癌物，并能损伤肝脏和肾脏。戴合适的手套和安全眼镜，并始终在化学通风橱中使用。

氯哇，通过吸入、摄入或经皮肤吸收，可能有害。长时间接触能导致永久性眼睛损伤。戴合适的手套和安全性护目镜。

$(CH_3)_2SO_4$，见硫酸二甲酯

柠檬酸，是一种刺激性物，通过吸入、摄入或经皮肤吸收，对人体有害。对眼睛构成严重损害的危险。穿戴合适的手套和护目镜，不要吸入粉尘。

考马斯亮蓝（CBR），通过吸入、摄入或经皮肤吸收，可能有害。穿戴合适的手套和安全眼镜。

DEAE，见二乙氨基乙醇

脱氧胆酸钠盐（DOC），通过吸入、摄入或经皮肤吸收，可能有害。不要吸入粉尘，戴合适的手套和安全眼镜。

二乙氨乙基（DEAE），通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜，并在化学通风橱中使用。

庚二亚氨酸二甲酯（DMP），对眼睛、皮肤、黏膜和上呼吸道有刺激。通过吸入、摄入或经皮肤吸收，可发挥有害效应。避免吸入雾剂，戴合适的手套、面罩和安全眼镜。

硫酸二甲酯（DMS），$(CH_3)_2SO_4$，极毒，是一种致癌剂，避免吸入雾剂，戴合适的手套和安全眼镜，仅在化学通风橱中使用。处理含有硫酸二甲酯的溶液时，将硫酸二甲酯缓慢倒入氢氧化钠或氢氧化铵溶液中，并使其在化学通风橱中静置过夜。在重返实验室清理泼洒之前，联系当地的安全办公室。

二硫苏糖醇（DTT），是一种散发恶臭气味的强还原剂。通过吸入、摄入或经皮肤吸收，可能有害。当对固体形式或高度浓缩的储液进行操作时，戴合适的手套和安全眼镜，并在化学通风橱中使用。

DMP，见庚二亚氨酸二甲酯

DMS，见硫酸二甲酯

二甲基硫醚 DOC，见去氧胆酸钠盐

干冰，见二氧化碳

DTT，见二硫苏糖醇

ENU，见 *N*-亚硝基-*N*-乙基脲和 *N*-乙基-*N*-亚硝基脲

乙醇（EtOH），CH_3CH_2OH，高度易燃，并且通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜，远离热、火星及明火。

乙醇胺，$HOCH_2CH_2NH_2$，具有毒性，通过吸入、摄入或经皮肤吸收，可能有害。小心操作，避免任何形式的皮肤接触。戴合适的手套和护目镜，并在化学通风橱中使用。乙醇胺具有强腐蚀性，并且与酸剧烈反应。

溴化乙锭，是强诱变剂，具有毒性，并且可能造成不可逆的效应。咨询当地公共机构的安全办公室，获得具体操作和处理程序。关于特定的操作和处理步骤，向当地的州级安全办公室咨询，避免吸入粉尘。当操作含有这种染色剂的溶液时，戴合适的手套。

乙酸乙酯，易燃，通过摄入可能致命。通过吸入或经皮肤吸收，也可能对人体有害，并导致延迟的靶器官影响。戴合适的手套和护目镜，不要吸入粉尘，在通风良好的地方使用。

乙二醇，是一种致畸剂，通过吸入、摄入或经皮肤吸收，可能有害，易燃。戴合适的手套和防护眼镜，并在化学通风柜中使用。

乙基亚硝基脲（ENU），见 *N*-亚硝基-*N*-乙基脲

EtOH，见乙醇

硫酸亚铁，$FeSO_4 \cdot 7H_2O$，通过吸入、摄入或经皮肤吸收，可能有害。穿戴合适的手套和安全眼镜。

$FeSO_4 \cdot 7H_2O$，见硫酸亚铁

FLAG-M2 树脂，见树脂

甲醛，HCHO，具有很强的毒性和挥发性，是一种可能的致癌物。很容易通过皮肤吸收，并刺激或破坏皮肤、眼睛、黏膜和上呼吸道。避免吸入蒸汽，戴合适的手套和安全眼镜，并始终在化学通风柜中使用，远离热源、火花和明火。

甲酰胺，致畸。蒸汽对眼睛、皮肤、黏膜和上呼吸道具有刺激作用。通过吸入、摄入或皮肤吸收，对人体有害。当使用甲酰胺浓缩液时，戴合适的手套和安全眼镜，并始终使用化学通风柜。尽量保持工作液密闭，以防蒸发。

盐酸胍，对黏膜、上呼吸道、皮肤和眼睛具有刺激作用。通过吸入、摄入或皮肤吸收，对人体有害。戴合适的手套和安全眼镜，避免吸入粉尘。

H_3BO_3，见硼酸

HCHO，见甲醛

HCl，见盐酸

H_3COH，见甲醇

H_2O_2，见过氧化氢

$HOCH_2CH_2NH_2$，见乙醇胺

$HOCH_2CH_2SH$，见β-巯基乙醇

盐酸，HCl，易挥发，通过吸入、摄入或经皮肤吸收，可能会有致命危险。它对黏膜、上呼吸道、眼睛和皮肤极有破坏性。戴合适的手套和安全眼镜，在化学通风橱中小心翼翼地使用。当对大量盐酸进行操作时戴护目镜。

过氧化氢，H_2O_2，对皮肤具有腐蚀性、毒性和极大的损伤性。通过吸入、摄入或经皮肤吸收，可能对人体有害。戴合适的手套和安全眼镜，并只在化学通风橱中使用。

咪唑，具有腐蚀性，通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜，并在化学通风柜中使用。

IPTG，见异丙基-β-D-硫代半乳糖苷

异丙醇，易燃，具有刺激性。通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜，不要吸入蒸汽，远离热源、火花和明火。

异丙醇（isopropyl alcohol），见异丙醇（isopropanol）

异丙基-β-D-硫代半乳糖苷（IPTG），通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜。

KCl，见氯化钾

$KMnO_4$，见高锰酸钾

KOH，见氢氧化钾

亮抑酶肽（亮肽素）（或其半硫酸盐），通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜，并在化学通风橱中使用。

液氮（LN_2），因极端温度可致严重损害。处理冷冻样品要特别小心。不要吸入蒸汽。液氮漏入冷冻小瓶后，一旦从液氮中移出，可致试管膨大。若可能，使用带O环密封圈的小瓶。戴低温手套和面罩，不要让液氮溅到你的衣服上，不要吸入蒸汽。

LN_2，见液氮

氯化镁，$MgCl_2$，一旦吸入、摄入或经皮肤吸收，则可能有害。戴合适的手套和安全眼镜，并在化学通风橱中使用。

硫酸镁，$MgSO_4$，一旦吸入、摄入或经皮肤吸收，则可能有害。戴合适的手套和安全眼镜，并在化学通风橱中使用。

MeOH或H_3COH，见甲醇

β-巯基乙醇（2-巯基乙醇），$HOCH_2CH_2SH$，一旦吸入或经皮肤吸收，可能会有致命危险，摄入后会有害。高浓度β-巯基乙醇对黏膜、上呼吸道、皮肤和眼睛具有极强的破坏性。戴合适的手套和安全眼镜，并始终在化学通风橱中使用。

$MgCl_2$，见氯化镁

$MgSO_4$，硫酸镁

NaOH，见氢氧化钠

$(NH_4)_2S_2O_8$，见过硫酸铵

Ni-NTA树脂，见树脂

N-亚硝基-*N*-乙烷基脲（ENU），可能的致癌物和强突变剂。通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜，并在化学通风中使用。不要吸入粉

尘。用 1 mol/L NaOH 净化 ENU 接触的所有材料。

Nonidet P40，导致严重的眼睛刺激和灼伤。通过吸入、摄入或经皮肤吸收，可能有害。穿戴合适的手套和护目镜。

α-^{32}P，见放射性物质

γ-^{32}P，见放射性物质

PEI，见聚乙烯亚胺

酚，具强腐蚀性，能导致严重灼伤。戴手套、护目镜，穿防护衣，并始终在化学通风橱中使用。用大量水冲洗酚接触的任何皮肤区域，用肥皂和水洗涤，不要使用乙醇洗。

苯甲基磺酸氟（PMSF），是具有强毒性的胆碱酯酶抑制剂。它对呼吸道黏膜、眼睛和皮肤极具破坏性。通过吸入、摄入或经皮肤吸收，可能会有致命危险。戴合适的手套和安全眼镜，并始终在化学通风橱中使用。一旦接触，立即用大量水冲洗眼镜和皮肤，丢弃被污染的衣服。

哌啶，具有强毒性，对眼睛、呼吸道和胃肠道有腐蚀性。和酸发生剧烈反应，使试剂氧化。通过吸入、摄入或经皮肤吸收，可能有害。不要吸入蒸汽，远离热源、火星和火焰，戴合适的手套和安全眼镜，并始终在化学通风橱中使用。

PMSF，见苯甲基磺酰氟

聚丙烯酰胺，被认为是无毒的，但应谨慎对待，因为它可能含有少量的未聚合的丙烯酰胺（见丙烯酰胺）。

聚乙烯亚胺（PEI），可能会引起眼睛和皮肤灼伤。通过吸入、摄入或经皮肤吸收，可能对人体有害。戴合适的手套和护目镜。

氯化钾，KCl，通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜。

氢氧化钾，KOH 及 KOH/甲醇，剧毒，如果吞食，可能是致命的。通过吸入、摄入或经皮肤吸收，可能有害。溶液有腐蚀性，并可造成严重灼伤。应该非常小心操作。戴合适的手套和护目镜。

高锰酸钾，$KMnO_4$，是一种刺激剂和强氧化剂。与有机物混合，可能形成爆炸性混合物。所有溶液在化学通风柜中使用，不要与盐酸混合。

蛋白 A-琼脂糖珠，见树脂

蛋白酶 K，是一种刺激物，通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜。

放射性物质：当规划一个涉及使用放射性的实验时，包括同位素的物理化学性质（半衰期、辐射类型和能量）、放射性的化学形态、放射性浓度（放射性比活度）、总数，及其化学浓度。订购和使用差不多真正需要的量。在操作放射性物质时，始终戴合适的手套、实验室工作服和护目镜。X 射线和 γ 射线是波长很短的电磁波，可由技术设备或放射性物质产生。它们可能从放射源上均质发出，或可能聚焦成束。它们潜在的危险取决于暴露的时间、经历的强度，以及使用的波长。请注意，恰当的屏蔽通常是铅或其他类似的材料。屏蔽的厚度是由 X 射线或 γ 射线的能量决定的。关于放射性材料的恰当

使用和处理，可向当地的安全办公室咨询，以获得进一步的指导。使用放射性同位元素后，始终要彻底监测。从下面的网站上可找到一个进行日常放射性计算的简便计算器：http://graphpad.com/quickcalcs。

树脂，易燃，并怀疑是致癌物。未聚合的组分和粉尘可能导致毒性反应，包括长期暴露形成的接触过敏。避免吸入蒸汽和粉尘，戴合适的手套和护目镜，并始终在化学通风柜中使用。对这些化学药品的敏感性可能随着反复接触而发展，远离热源、火花和明火。

核糖核酸酶 A 是一种刺激物，通过吸入、摄入或经皮肤吸收，对人体有害。戴合适的手套和安全眼镜，不要吸入粉尘。

SDS，参见十二烷基硫酸钠

二甲砷钠，可能是致癌物，并含有砷。通过吸入、摄入或经皮肤吸收，可能有致命危险，也可能对未出生婴儿有害。接触或吸入的效应可能是迟发的。不要吸入粉尘，戴合适的手套和安全护目镜，并在化学通风橱中使用。

柠檬酸钠，见柠檬酸

十二烷基硫酸钠（SDS），具有毒性，是一种刺激物，并对眼睛具有造成严重损害的危险。通过吸入、摄入或经皮肤吸收，对人体有害。戴合适的手套和护目镜。不要吸入粉尘。

氢氧化钠，NaOH 及含有 NaOH 的溶液，具有强毒行和腐蚀性，应小心操作。戴合适的手套和面罩。所有浓缩碱应以同样的方式处理。

Source 15Q 树脂，见树脂

亚精胺，可能具有腐蚀性，并造成眼睛和皮肤灼伤。通过吸入、摄入或经皮肤吸收，对人体有害。损伤效应可能会推迟。戴合适的手套和护目镜，并在化学通风柜中使用。

精胺，具有腐蚀性，并且可能造成眼睛和皮肤的灼伤。通过吸入、摄入或经皮肤吸收，可能有害。损伤效应可能会推迟。戴合适的手套和护目镜，并在化学通风柜中使用。

TEMED，见 *N*，*N*，*N* ′，*N*′-四甲基乙二胺

N，*N*，*N*′，*N*′-四甲基乙二胺（TEMED），对眼睛和黏膜具有腐蚀性，并且通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套，并紧紧密封护目镜。

硫脲，可能是致癌物，通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜，并在化学通风橱中使用。

Tris，是一种刺激物，并通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜。

Tris 碱，是一种刺激物，并通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜。

Triton X-100，造成严重的眼睛刺激和灼伤。通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和护目镜，不要吸入蒸汽。

台盼蓝，可能是一种致癌物，通过吸入、摄入或经皮肤吸收，可能有害。不要吸入

粉尘，戴合适的手套和安全眼镜。

尿素，通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和安全眼镜。

UV 光源和（或）紫外辐射，有危险，并能损伤视网膜。绝不要用裸眼看未经遮挡的 UV 光源。实验室中常见的紫外光源包括手提式紫外灯和投射仪。仅可通过能够吸收有害波长光线的滤光片或安全防护镜观察。UV 射线（参见 UV 光源）危险，特别是对于眼睛。UV 射线还具有致诱变性和致癌性。紫外辐射也具有致突变性和致癌性，尽量减少暴露，确保紫外光源充分遮挡。当在 UV 光源下对材料进行操作时，戴合适的保护性手套。

二甲苯，易燃，高浓度下可能具有麻醉作用。通过吸入、摄入或经皮肤吸收，可能有害。戴合适的手套和防护眼镜，并在化学通风柜中使用，远离热源、火花和明火。

二甲苯青，见二甲苯

索　引

(Q-2933.31)

生命科学实验指南系列·典藏版





科学出版中心 生物分社
联系电话：010-64012501
E-mail：lifescience@mail.sciencep.com
网址：http://www.lifescience.com.cn
销售分类建议：分子生物技术



赛拉艾芙
生命科学订阅号


ISBN 978-7-03-047486-5
9 787030 474865 >

定价（全套）：4500.00元